Toward a meta-methodology for real-world problem solving

by

Ryan Reed

Dissertation presented for the degree of
Doctor of Philosophy in Industrial Engineering
in the Faculty of Engineering at Stellenbosch University

Promoter: Prof JH van Vuuren

March 2020

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 2020

Abstract

There are complexities inherent to the subjective experience of life that render real-world problem solving exceedingly subtle and difficult. Yet, all too often, it is not the complex nature of problems that result in their persistence, but rather the ineffective fashion in which practitioners dedicated to their resolution respond thereto.

The process of real-world problem solving comprises three parts: The development of a formal formulation of a problem considered, the identification of an adequate response thereto, often by means of analytical modelling, and finally, the practical implementation of that response. In the history of real-world problem solving, the first of these phases has often been neglected and, as a result, solution methodologies dedicated to the formulation of problems, complicated by the dynamics of human nature, have failed to address several phenomena relevant to the success of problem resolution. Accordingly, the work documented in this dissertation is dedicated first to the identification of such phenomena, and then to the development and application of a meta-methodology dedicated to their management.

The phenomena addressed include the identification of stakeholders relevant to a given instance of problem solving, the dissimilar states in which real-world systems exist and the effect thereof upon the resolution of problems therein, the sociological paradigm-based limitations of a select group of problem formulation methods, and the mitigation of irrational influences on human decision making. In an attempt to uncover dynamics relevant to the formulation of real-world problems hidden from a purely qualitative analysis, the unique utility of mathematical modelling is additionally employed to describe the context for which the proposed meta-methodology is designed.

In order to illustrate the practical applicability of the meta-methodology proposed in this dissertation, several of its components are applied, first in a hypothetical fashion, to the context of the South African education system during the period of the *fees must fall* movement and then in a practical fashion, to the context of the South African energy sector.

Uittreksel

Daar is kompleksiteite inherent aan die subjektiewe ervaring van die lewe wat werklike probleemoplossing uiters subtiel en moeilik maak. Tog is dit al te dikwels nie die komplekse aard van probleme wat hul volharding veroorsaak nie, maar eerder die ondoeltreffende manier waarop praktisyns toegewy aan die oplossing daarvan, daarop reageer.

Die proses van werklike probleemoplossing bestaan uit drie dele: Die ontwikkeling van 'n formele formulering van die probleem wat oorweeg word, die identifisering van 'n gepaste respons daarvoor, dikwels deur analitiese modellering, en laastens die praktiese implementering van die respons. In die geskiedenis van werklike probleemoplossing is die eerste van hierdie fases al te dikwels verwaarloos, en as gevolg daarvan het oplossingsmetodologieë vir die formulering van probleme, gekompliseer deur menslike natuurlike dinamika, soms nie daarin geslaag om verskynsels aan te spreek wat relevant is tot die sukses van die probleemoplossing nie. Gevolglik word die werk wat in hierdie proefskrif gedokumenteer is, eers op die identifikasie van sulke verskynsels toegespits, en daarna aan die ontwikkeling en toepassing van 'n meta-metodologie vir hul bestuur toegewy.

Die verskynsels wat aangespreek word, sluit in die identifisering van belanghebbendes wat relevant is vir 'n gegewe probleemoplossing, die uiteenlopende toestande waarin werklike wêreldstelsels bestaan en die effek daarvan op die oplossing van probleme daarin, die sosiologiese paradigma-gebaseerde beperkings van 'n bepaalde groep probleem formuleringsmetodes, en die teenwerking van irrasionele invloede op menslike besluitneming. In 'n poging om dinamika te identifiseer wat relevant is vir die formulering van werklike probleme wat in 'n suiwer kwalitatiewe analise verborge bly, word die unieke nut van wiskundige modellering ook ingespan om die konteks waarvoor die voorgestelde meta-metodologie ontwerp is, te beskryf.

Ten einde die praktiese toepaslikheid van die meta-metodologie wat in hierdie proefskrif voorgestel word, te illustreer, word verskeie komponente daarvan eerstens op hipotetiese wyse in die konteks van die Suid-Afrikaanse onderwysstelsel gedurende die periode van die *fees must fall*-beweging en daarna prakties in die konteks van die Suid-Afrikaanse energiesektor toegepas.

Acknowledgements

The author wishes to acknowledge the following people and institutions for their various contributions towards the completion of this work:

- My supervisor, Professor JH van Vuuren, for granting me the opportunity to wander the road less travelled, and for his wise counsel along the way.
- The members of the *Stellenbosch Unit for Operations Research in Engineering* (SUnORE), for making the last two years a banter-filled experience that will be remembered with fondness.
- The SUnORE research group, for financial support in the form of a post-graduate bursary.
- The guy with the keen eye, Benedikt Meylahn, for his first-class perusal.
- The squad of 57 Jonkershoek Street, for teaching me that electronic engineers do actually dance (even if it is in jeans and tackies), that artists do, in fact, care about the state of one's living quarters, and that blossoms do indeed bloom in the spring.
- My family. Our story is needled into the very fabric of this dissertation; may spring come soon.

Table of Contents

List of Acronyms

ANC: African National Congress

ASP: Association of Spirit Producers

CIA: Central Intelligence Agency

CSH: Critical systems heuristics

CST: Critical systems thinking

LSI: Local systems intervention

NIA: National Institute for Alcohol

NPHI: National Public Health Institute

SAST: Strategic assumption surfacing and testing

SCA: Strategic choice approach

SODA: Strategic options development and analysis

SRC: Student Representative Counsel

SSEG: Small-scale-embedded-generation

SSM: Soft systems methodology

SSS: Social systems sciences

TRSDP: Transport and Road Safety Division of the Police

TSI: Total systems intervention

List of Figures

List of Tables

CHAPTER 1

Introduction

Contents

1.1 Background

It has long been recognised that the subjective experience of everyday life is exceptionally complex. Likened to the current of a turbulent stream, humanity often finds itself carried along by a dynamic flux of events, ideas, relationships, emotions, and actions — all mediated through the slippery agency of communication and all continually changing. Simply put, the natural state of the world constitutes a messy state of affairs.

Despite the obvious truth of the above statement, history testifies that practitioners dedicated to the solution of real-world problems were surprisingly dilatory in recognising both the importance and the difficulty of formulating such problems in a manner that actually facilitates their effective resolution. In fact, the intricacies of problem formulation only entered the fray of academic and professional debate some time during the mid-20th century [234]. As the discipline of *systems thinking* constitutes the theoretical foundation upon which the aforementioned group of practitioners often based the resolution of problems, much can be learnt about the manner in which the consequential nature of problem formulation entered the realm of conscious awareness, by delineating the paradigms of enquiry that have dominated systems thinking throughout the course of its historical development.

The paradigm of enquiry that "put systems thinking on the map," so to speak, has come to be known by the generic name *hard systems thinking* — in essence, a paradigm that prioritises the notions of rationality, efficient means and control. The first documented instance of hard systems thinking is often attributed to the period during and immediately after World War II, constituting a time of crisis during which military planners recognised the utility of scientific methods in supporting wartime decision making [233]. Significantly, the majority of the methods designed and applied during this period were based upon the principles of hard systems thinking, and furthermore, were employed with considerable success. As a result, the hard systems school

of thought became the dominant paradigm of enquiry associated with system theorists and real-world practitioners alike during the decades that followed.

Interestingly, it took approximately 15 years of practical application and academic scrutiny, before the limitations of hard systems thinking came to be recognised [234]. To elucidate, during the mid-20th century the methods of hard systems thinking slowly made their way into the volatile context of corporate endeavour and, as a result, met their match. One of the fundamental assumptions associated with the hard systems school of thought is that problems are naturally inclined to their formulation and, as such, solution methodologies based thereupon hardly ever include considerations of the process of problem formulation [99]. In the punctilious context of its inception, this lack of attention may have been inconsequential, but not so in the more human-centric context of corporate endeavour. The plethora of human-induced dynamics often relevant to the solution of corporate problems, seems to have rendered such problems of the kind for which a solution approach based solely on the analytical methods of hard systems thinking is rather ill-suited [42, 475].

Ironically, in the history of real-world problem solving, it was thus the failures of hard systems thinking (a paradigm based upon the assumption that problem formulation is superfluous) that firmly established (at least in the mind of several theorists) the need to address the difficulties that surround the formulation of real-world problems. Accordingly, several theorists of the time (including renowned academics such as Churchman and Ackoff [232]) elected to distance themselves from the conceptions of hard systems thinking, and proceeded to develop a paradigm of enquiry based upon the general principles of systems theory that is capable (according to them) of accommodating the noted, nuanced complexities related to the formulation of real-world problems. Significantly, the development of this paradigm, which came to be known as *soft systems thinking*, became for the notion of problem formulation what one may describe as a place to call home in the realm of academic endeavour.

In view of the aforementioned manner in which both the difficulty, and the importance of problem formulation historically entered into the realm of conscious awareness, the obvious question of interest is simply: Did such awareness actually result in the development of practical solution methodologies capable of accommodating the noted difficulties in formulating real-world problems of the kind saturated by the dynamics of human nature? The answer to this question is both *yes*, and *no*. Solution methodologies dedicated to the formulation of real-world problems have certainly been developed [233], but as Jackson delineates in his book *Systems methodology for the management sciences* [232], the field of problem formulation has by no means matured to the point where its methods are indeed capable, in general, of formulating problems in a manner that truly facilitates their resolution [234]. Accordingly, the work documented in this dissertation is an attempt to further the field of problem formulation some way towards its maturity.

1.2 Problem statement

Solution methodologies dedicated to the formulation of real-world problems of the kind complicated by the dynamics of human nature often fail to address several phenomena relevant to the success of their application. Accordingly, the aim in this dissertation is first to identify such phenomena, and then to develop, and demonstrate the application of, a meta-methodology dedicated to their management, which may be employed in conjunction with existing methodologies that are designed to facilitate the initial phases of real-world problem solving.

1.3 Dissertation objectives

The following ten objectives are pursued in this dissertation:

I To *conduct* a thorough survey of the academic literature related to:

(a) human decision making under conditions of uncertainty,

(b) human decision making in the context of social stimuli,

(c) common approaches to debiasing,

(d) paradigms of thought in systems theory,

(e) problem solution methodologies,

(f) typologies for the classification of real-world problems, and

(g) stakeholder selection theory.

II To *develop* a generic method for identifying stakeholders relevant to a certain instance of real-world problem solving.

III To *delineate* the dissimilar states in which real-world systems exist, the mechanisms by which these states come about, and the strategies according to which they may be dealt with.

IV To *propose*, based upon the work conducted in pursuit of Objective III, a method designed to support practitioners in minimising the cost of real-world systems in chaos.

V To *present*, based upon the literature review of Objectives I(c)–(f), an analysis of the problem structuring methods that exist in system literature, toward supporting their case-specific selection.

VI To *design*, based upon the literature review of Objectives I(a)–(c), a decision-making checklist aimed at mitigating the irrational influence on human decision making.

VII To *formulate* a meta-methodology for the formulation of real-world problems, based on the work carried out in pursuit of Objectives II–VI.

VIII To *model* mathematically the competitive interaction of entities relevant to the context of use for which the meta-methodology of Objective VII is designed, in an attempt to uncover the mechanisms that govern such interaction, hidden from purely a contemplative analysis.

IX To *illustrate* the practical applicability of the methods proposed in pursuit of Objectives II, IV, and VI by means of either their theoretical application, their practical application, or the scrutiny of an expert analyst.

X To recommend sensible follow-up work related to the work in this dissertation which may be pursued in future.

1.4 Scope delimitation

The intricate spectrum of phenomena relevant to the resolution of real-world problems is such that the pursuit of their delineation has remained elusive throughout the course of scientific

enquiry. Accordingly, it should come as no surprise that the meta-methodology proposed in this dissertation is designed to address only a select portion of the phenomena relevant to the solution of real-world problems. In order to elucidate this class of phenomena considered, consider Figure 1.1, depicting an oversimplified, high-level overview of the activities relevant to the process of real-world problem solving, namely *problem identification*, *problem formulation*, *response formulation* and *response implementation*[1]. In view of the figure, note that the meta-methodology of Objective VII, and the respective components that constitute it as outlined in Objectives II–VI, relate only to the process of problem formulation and, more specifically, only to aspects thereof not adequately addressed in existing methodologies dedicated thereto, as determined later during the literature review in pursuit of Objective I.

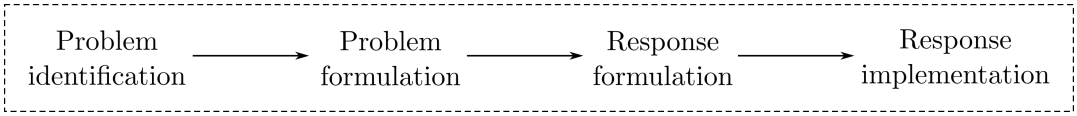


FIGURE 1.1: *A high-level overview of the process of real-world problem solving.*

In addition, the work presented in this dissertation is also limited to a certain context of application. More specifically, Objectives II–X are pursued only in the context of real-world systems that exhibit the following property: The entities contained therein are forced to collaborate despite pursuing conflicting outcomes. Note that this kind of system may exist for one of several reasons, including, but not limited to, the following: If multiple entities of the system are fundamentally linked to its purpose, if legislation governing such a system prohibits the exclusion of certain entities, or if no entity contained therein is capable of forcefully (and lawfully) extinguishing the presence of another.

In view of the above scoping delimitations, the work presented in this dissertation is adequately described as an attempt to support analysts during the initial phases of real-world problem solving purely by supplementing a selection of pre-existing solution methodologies dedicated to the formulation of problems.

1.5 Dissertation organisation

In addition to this introductory chapter and the one that follows, which is dedicated to a discussion on the research methodology adopted, the work presented in this dissertation is organised into three parts, collectively comprising a further thirteen chapters. The first part, Part I, contains six chapters and is dedicated to a review of the academic literature related to matters relevant to the topic of this dissertation, as described in §1.3. To elucidate, Part I opens in Chapter 3 with a brief introduction into the mechanisms of human cognition in §3.1, and this is followed by a discussion on the intricacies of judgement heuristics in §3.2. This sets the stage for an account of the cognitive biases that accompany the application of such heuristics in §3.4–3.8, after which Chapter 3 comes to a close with a short summary in §3.9.

The next chapter, Chapter 4, is dedicated to an enquiry into the social nature of human decision making. After this topic is briefly introduced in §4.1 with a discussion on the paradoxical nature of group behaviour, the social dynamics related to decision making in the presence of relevant others is explicated in §4.2–§4.4. The chapter then concludes with a brief summary in §4.5.

[1]Problem formulation represents the process of developing a formal representation of a problematic situation that adequately facilitates the pursuit of an appropriate response to the problem considered, while response formulation represents the process of actually developing an appropriate response by means of analytical modelling.

In accordance with the objectives pursued in this dissertation, Chapter 5 is dedicated to an enquiry into the field of debiasing. The chapter opens with a brief introduction to the field in §5.1, after which the schools of thought that underlie the methods employed therein are reviewed in §5.2–§5.4. After a collection of methods based upon the reviewed schools of thought has been reviewed in §5.5, the chapter comes to a close with a short summary in §5.6.

Chapter 6 embodies quite a shift in the focus of discussion, opening in §6.1 with an account of the origins of systems theory. Excluding a short summary of matters discussed in §6.10, the remainder of the chapter is dedicated to a review of the paradigms that dominate systems theory, as well as the respective solution methodologies related thereto, in §6.1–§6.9.

Chapter 7 opens with a brief introduction of the notion of dissimilar problem types in §7.1, which is followed by a description of three unique problem typologies for the classification of real-world problems in §7.2, §7.3 and §7.4, respectively. The chapter finally closes with a short summary in §7.5.

Chapter 8, with which Part I concludes, is dedicated to a review of the methods that dominate the field of stakeholder selection theory. The chapter opens in §8.1 with a review of several methods that facilitate the identification of stakeholders, and this is followed in §8.2–§8.5 by an analysis of a collection of methods considered to facilitate the classification of stakeholder, and in §8.6 by a discussion of a collection of methods that facilitates an investigation into the network of relationships among stakeholders. In §8.7, a high-level comparison of the methods discussed is presented, and the chapter is thereafter brought to a close with a brief summary in §8.8.

Part II is the heart of the dissertation and consists of a further five chapters, starting in Chapter 9 with a delineation of the novel meta-methodology proposed. Accordingly, the chapter opens with a brief overview of the phases that constitute the proposed meta-methodology in §9.1, setting the stage for an in-depth discussion on each of its phases in §9.2 and §9.3. The chapter concludes with a concise summary of matters discussed in §9.4.

Chapter 10 contains an account of the first stage of evaluation of the proposed meta-methodology. This involved its hypothetical application in the context of a theoretical case study in retrospect. The chapter commences with an account of the selected context of application in §10.1, and this is then followed by the hypothetical application of the proposed meta-methodology in that particular context in §10.2. Thereafter, an additional phase in the meta-methodology, designed to address a potentially detrimental phenomenon uncovered during its hypothetical application, is introduced in §10.3. The chapter finally comes to a close with a short summary in §10.4.

In Chapter 11 the proposed meta-methodology is subjected to the critique of an expert analyst. The chapter opens with a brief account of the feedback received from the analyst, in §11.1, and this is followed by the presentation of an updated version of the meta-methodology in §11.2. The chapter then comes to a close with a short summary, in §11.3.

Chapter 12 is dedicated to the documentation of the final phase of evaluation of the proposed meta-methodology — its application to a practical case study. Accordingly, the chapter commences with a brief description of the context in which the meta-methodology is applied in §12.1, and this is followed with an account of its actual application in §12.2 and §12.3. The chapter finally comes to a close with a brief summary, in §12.4.

In the final chapter of Part II, the subject matter becomes slightly more mathematical. More specifically, Chapter 13 comprises an analytical model developed to describe (to a certain extent) the competitive interaction among the entities relevant to the context of application for which the proposed meta-methodology is designed. The derivation of the model commences with

the introduction of a governing assumption in §13.1, and this is followed by a mathematical description of the relationship between several elements of interest in §13.2–§13.3. The notion of infeasibility is briefly discussed in §13.4, after which an analysis of the proposed model is conducted in §13.5. Next, an extension to the model is presented in §13.6 and §13.7, and this is followed in §13.8 by the generalisation of a result related to the manner in which entities compete in an endeavour, depending on the properties that characterise it. The chapter is finally brought to a close with a short summary in §13.9.

The final part of this dissertation, Part III, consists of two concluding chapters. The first, Chapter 14, contains a summary of, and a reflection on, the collective body of work presented in the dissertation, and the second, Chapter 15, comprises a suggestion of potential avenues of further investigation that may be pursued as future work.

CHAPTER 2

Research methodology

Contents

This chapter is dedicated to a discussion on the particular research paradigm and research methodology adopted in this dissertation. Importantly, the particular selections made in this regard were subject to, and approved by, an expert in such matters[1].

The chapter opens in §2.1 with a discussion on the theoretical vantage point adopted, and this is followed in §2.2 with an brief overview of the research methodology utilised. In §2.3, another expert analyst is introduced whose opinions are later considered during a validation process followed in respect of the meta-methodology put forward, and then the chapter is brought to a close with a brief summary in §2.4.

2.1 The research paradigm

During the initial phases of a research project, a series of decisions must be made pertaining to the paradigm of enquiry adopted. The first such decision, dedicated to the selection of a theoretical vantage point from which phenomena are to be studied, entails delineating the particular *ontological* and *epistemological* points of view that govern the process of acquiring knowledge. Individually, these elements address the assumed state of reality, and the nature of knowledge, respectively. Collectively, they embody what is here called a *research paradigm*. Several such paradigms exist in the academic literature [196, 397, 428], but two appear dominant, namely the *positivist-objectivist* paradigm and the *interpretivist-constructivist* paradigm [397]. Before specifying which of these paradigms are, in fact, adopted to govern the work documented in this dissertation, it is beneficial to briefly delineate the qualities that characterise them.

Perhaps the most defining property of the positivist-objectivist paradigm is that it assumes independence between the predefined perceptions of a researcher and properties of a studied phenomenon. That is, objective, value-free research is considered possible [428]. In its view of

[1] The expert in question is Professor SS Grobbelaar from the Department of Industrial Engineering at Stellenbosch University [195]

knowledge, meaning is deemed independent of observation, in the sense that the recognition of a particular phenomenon does not suddenly grant it meaning — rather, the process of discovery simply represents the detection of an objective, hidden truth that was there all along. On the opposite end of the spectrum, one finds a paradigm in which knowledge of the world is considered influenced by, and dependent upon, the particular perspectives of the observer, which, in turn, are considered the product of both reflected upon and lived experiences — the interpretivist-constructivist paradigm [353]. In contrast with positivist-objectivist tradition, this paradigm holds that insight is the product of exploring (not discovering) the social world of individuals from the subjective vantage point of the observer. The research process associated with the paradigm in question is, furthermore, largely *inductive* in the sense that its purpose is to generate theory from data rather than using data to test theory [397]. Finally, in the interpretivist-constructivist paradigm, the predefined perceptions of a researcher and the nature of what is observed are viewed to be interdependent such that totally objective, value-free research is considered impossible.

In accordance with the above discussion, and the nature of the work conducted in this dissertation, the interpretivist-constructivist paradigm is selected as the governing paradigm of enquiry.

2.2 The research methodology

Following the establishment of a particular research paradigm, an appropriate research methodology has to be selected. To that end, a design framework formalised by Herman [211] in 2018 for aiding researchers in the development of multidisciplinary solution methodologies is adopted to guide the research conducted in this dissertation. Notably, the framework in question is an adaption of the well-known *conceptual framework analysis* procedure proposed earlier by Jabareen [228] and, as may be seen in Figure 2.1, consists of four main parts, namely *research the design aims and make sense of the literature*, *formulate a preliminary framework*, *evaluate and adapt the framework*, and finally, *finalise the framework and management tool*.

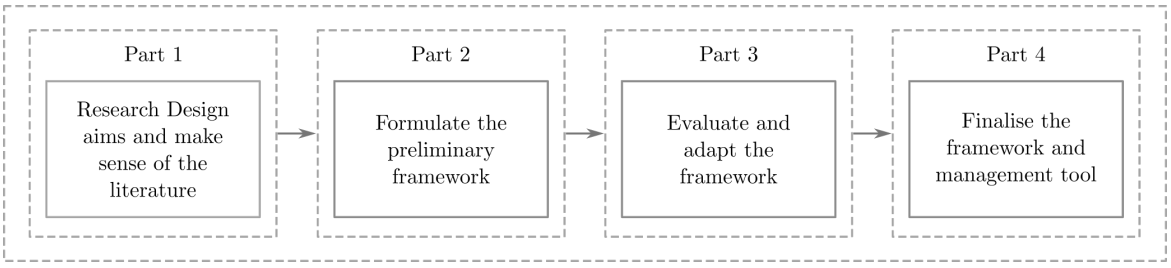


FIGURE 2.1: *A design framework for facilitating the development of multidisciplinary solution methodologies [211].*

Part 1 of the above framework is designed to aid analysts in exploring the research landscape and academic literature related to a topic of enquiry, and is separated into four distinct phases, namely *describe the research question*, *specify the research aims*, *conduct a background literature review*, and *conduct a structured literature review*, as depicted in Figure 2.2. As the labels suggest, these phases involve formalising the research question underlying the development of a particular methodology, describing the specific objectives that are pursued in pursuit of addressing the stated research question, reviewing the academic literature in order to filter relevant from irrelevant documentation, and finally, conducting an in-depth review of the literature deemed

relevant[2]. Excluding the third phase of this process, which is alternatively described as an informal, undocumented review of the academic literature, the execution of these phases relate to the work documented in §1.2, §1.3, and Chapters 3, 4, 6 and 7 of this dissertation, respectively.

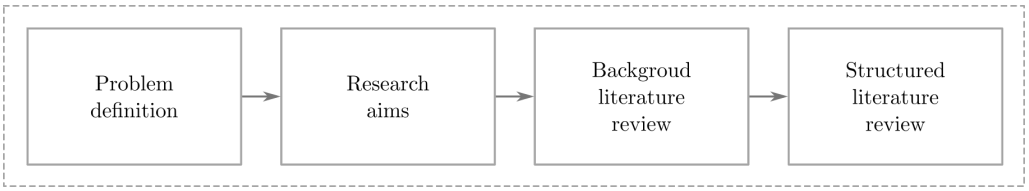


FIGURE 2.2: *A schematic depiction of the constituent phases of Part 1 of the design framework adopted, as adapted from Herman [211].*

Part 2 of the framework consists of four components that are collectively designed to facilitate an analysis of existing solution methodologies, similar to the one proposed, such that the insights of past researchers may be drawn upon to guide the development of the methodology proposed. As depicted in Figure 2.3, the constituent phases of Part 2 are *conduct focused conceptual literature review*, *identify existing methodologies and models*, *evaluate and assess current methodologies and models*, and *develop preliminary methodology*. Once again, some of these phases are conducted in an undocumented fashion as they simply inform the execution of the phases that proceed them. More specifically, only the execution of the third and fourth phases are presented in this dissertation in Chapters 5 and 8, and in Chapter 9, respectively.

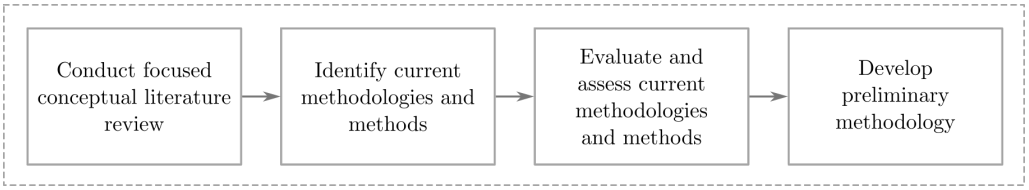


FIGURE 2.3: *A schematic depiction of the constituent phases of Part 2 of the design framework adopted, as adapted from Herman [211].*

Researchers often adopt the term validation[3] when discussing the regimen of review adopted in a particular study. Due to the nature of the work documented in this dissertation, however, its validation in the traditional sense of evaluating the utility thereof with several benchmark case studies in the literature is not possible as no such benchmarks exist. For this reason, the design framework proposed by Herman [211] is particularly well-suited to the work carried out in this dissertation. That is, Part 3 of the framework in question involves a multistage evaluation (instead of validation) process that facilitates the review and refinement of the instrument of analysis proposed [211]. As depicted in Figure 2.4, this process consists of five phases, namely *conduct a theoretical case study*, *modify the methodology*, *conduct interviews*, *modify the methodology*, and *conduct a practical case study*. In the order listed, these phases entail the hypothetical application of the proposed methodology to a theoretical case study in retrospect with the purpose of identifying shortcomings therein, presenting an updated methodology modified according to the shortcomings uncovered, subjecting the methodology to the critique of an expert analyst,

[2]In the context of this dissertation, the in-depth literature review is executed in accordance with the guidelines provided by Lecy and Beatty [281] in an article titled *Structured literature reviews using constrained snowball sampling and citation network analysis.*

[3]The term validation is considered by Herman [211] to represent the process by which a researcher attempts to prove the utility and accuracy of a contribution made.

presenting a second updated version of the methodology, and finally, the applying the updated methodology to a practical case study in real time. The work conducted in pursuit of these tasks is documented in Chapters 10, 11 and 12.

FIGURE 2.4: *A schematic depiction of the constituent phases of Part 3 of the design framework adopted, as adapted from Herman [211].*

Before discussing the final part of the framework proposed by Herman [211], there is a technicality related to the evaluation of the proposed methodology that should be noted. That is, any shortcoming identified during a particular phase of evaluation is to be subject only to the remainder of the phases of evaluation and only if the particular addition to, or component of, the methodology is, in fact, applicable to the phase of evaluation in question. With this in mind, consider the constituent phases of the final part of the design framework adopted in this dissertation, as depicted in Figure 2.5. As may be seen in the figure, Part 4 consists of three phases, namely *modified methodology*, *management tool*, and finally, *future work*. In essence, the former phases entail taking the most up-to-date version of the methodology proposed and transforming it into a form that facilitates its use in the hands of future practitioners, while the latter simply entails the description of potential improvements to the methodology that may be pursued in future. Notably, the final version of the methodology presented in fulfilment of Part 3 of the design framework is formulated with the requirements of Part 4 in mind such that the final version of the meta-methodology proposed in this dissertation may be found in Chapter 11, while potential improvements thereto, which may be pursued in future, are documented in Chapter 15.

FIGURE 2.5: *A schematic depiction of the constituent phases of Part 4 of the design framework adopted, as adapted from Herman [211].*

2.3 An expert analyst

In the introduction to this chapter it was made salient that the research methodology adopted to guide the work presented in this dissertation has been approved by an expert in such matters. The expert referred to then, and the expert referred to in Part 3 of the research methodology discussed above, are, however, two different individuals. Appropriately, the former is an expert in designing and conducting research, and the latter, in the actual application of solution methodologies of the kind proposed in this dissertation. Mr Brink van der Spuy is the second expert, is a seasoned consultant in applications such as strategy development, change management and corporate mediation, and has worked with many of the largest corporations in South Africa [73]. Mr van der Spuy is, furthermore, considered a particularly well-suited expert, given

his experience as a lecturer of business strategy and innovation at several esteemed universities, including Stellenbosch University and the University of South Africa [73].

2.4 Chapter summary

This chapter was dedicated to an account of the theoretical paradigm of enquiry selected to govern the research documented in this dissertation. Accordingly, the chapter opened in §2.1 with a brief discussion on the particular ontological and epistemological points of view assumed. Thereafter, a design framework, utilised as research methodology, was described in §2.2. Finally, an expert analysts, subsequently utilised to critique the meta-methodology proposed, was introduced in §2.3.

Part I

Literature Study

CHAPTER 3

The bounds of human rationality

Contents

"Real knowledge is to know the extent of one's ignorance." — Confucius

This chapter is devoted to a review of topics in the academic literature that are relevant to judgement and decision making under conditions of uncertainty. These topics are introduced in §3.1 by illustrating the dynamics of intuitive and analytical reasoning. Thereafter, several judgement heuristics that prevail upon the blurred boundary between intuitive and analytical reasoning are reviewed in §3.2. Finally, a number of cognitive biases that render judgement heuristics imperfect are considered in §3.3–§3.8, before the chapter closes with a brief summary in §3.9.

3.1 Intuition or reason

In his Nobel Prize-winning work, Herbert Simon [301, 421] argued that individual judgement is bounded in its rationality. Although this is a perfectly appropriate statement to the thinkers of the 21st century, the idea was rather preposterous to the *homo economicus* traditionalists of the day. The belief that individual judgement is based upon the rational processing of information toward the optimal pursuit of a subjectively defined end was at the core of decision theory when Simon formulated his hypothesis. Yet, his words stood the test of time: In an information-rich world, the wealth of information signifies the dearth of something else, that thing which information consumes — the attention of its recipients [301]. That is, a wealth of information

creates a poverty of attention and the need to allocate that attention appropriately, a task that signifies the first of several steps away from the path of rationality.

Simon's work encouraged researchers to pursue an understanding of actual decision making — rather than only focusing on the study of decision making within the vacuum of rationality. In time, these endeavours inspired two schools of thought within decision theory, namely *prescriptive decision theory*, and *descriptive decision theory* [1, 426]. Prescriptive models typically assume ideal circumstances in order to accommodate the best (*i.e.* most rational) path a person can take so as to make the most suitable decision in a specific context. In contrast, descriptive models of decision making are designed to define how people actually make decisions under real-world circumstances, regardless of their rationality.

The value of prescriptive models remains largely theoretical — people rarely, if ever, operate under the presumed circumstances of such theories [186]. In contrast, descriptive models are of great practical value. Rather than idealising the conditions of a typical decision making process, scholars are able to study and better understand the internal schemas people actually use when making decisions [51]. Daniel Kahneman [262, 456], one of the most respected decision theorists of our time, received a Nobel Prize in part for his work in descriptive theory. A topic of special interest to Kahneman was the distinction between intuition and reasoning [258]. Kahneman's work, among those of many others (see, for example, Epstein [145], Hammond [201], and Jacoby [235]), led to the establishment of a formal typology of cognition. Stanovich and West [432] were first to suggest the terms *System 1 thinking* (*i.e.* intuitive) and *System 2 thinking* (*i.e.* analytical) as descriptors of the typology.

System 1 thinking refers to our intuitive system, which is typically *effortless*, *fast*, *automatic*, *implicit* and *emotional* [44, 51, 215]. Consider Figures 3.1 and 3.2 as illustrations of the working of System 1 thinking.

FIGURE 3.1: *An illustration of System 1 thinking [51].*

Studying Figure 3.1, most analysts are of the opinion (and correctly so) that the second tree from the left is the tallest. This reflex judgement is a result of System 1 thinking and is of great practical value [51]. Imagine an employee tasked with the responsibility of identifying the tallest ten percent of trees in a large plantation. It would be laborious and inefficient to measure the height of every tree. According to System 1 thinking, some form of intuitive judgement should instead be utilised to save time and reduce effort [111]. In view of its efficiency it is tempting to think of System 1 thinking as a robust approach. Consider, however, the second example of System 1 thinking illustrated in Figure 3.2. After brief consideration, most analysts would agree that the surface of the table on the right could be described as being more square than that of the table on the left. This reflex judgement is once again a result of System 1 thinking, but in this instance it is, in fact, failing the analyst. To illustrate this failure, attempt the following System 2 strategy: Place a sheet of paper over the drawing and trace the top of either

table. Align the traced square over the other table and see how the tables align perfectly. The two illustrations above qualify System 1 thinking as an invaluable tool in the decision makers toolbox, but also one that should be used with care [51].

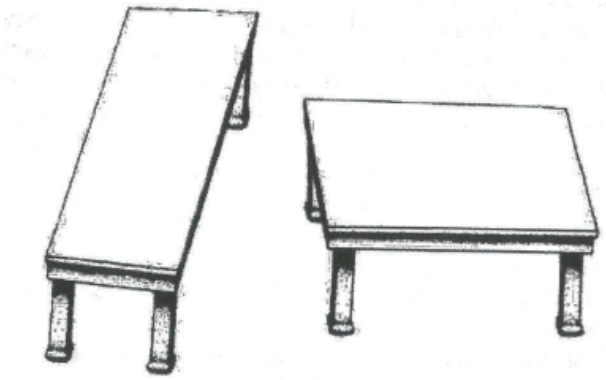


FIGURE 3.2: *An illustration of the limitations of System 1 thinking [51].*

System 2 thinking refers to one's analytical system working in tandem with the brain's *central working memory system*, a partnership appropriately described by the words *slow*, *conscious*, *laborious*, *explicit*, *logical*, and *sequential in nature* [200, 259]. Despite its limited capacity and slower speed of operation, System 2 thinking permits an analyst to engage in *abstract hypothetical thinking*, something that cannot be achieved by System 1 thinking [149]. Consider the case of a vehicle driver, for example, who has to select the appropriate gear during the execution of a left-turn manoeuvre. The driver could make his or her selection on the basis of past experience [12] — an intuitive process that requires little reflection. The driver may, however, also make the decision by constructing mental models of possible future events, hypothetically testing the effectiveness of several courses of action (*i.e.* employ abstract hypothetical thinking) [149]. Although one could definitely argue that simply relying upon past experience would be the most effective strategy in many instances of decision making in which time is of the essence, the distinctly human facility of simulating possible future events is of great importance — we cannot afford to learn by experience to avoid disasters such as nuclear wars, for example.

In view of the literature reviewed above, it is tempting to argue that the bounded nature of human decision making is limited to the mechanisms of System 1 thinking. In reality, however, this is not the case: System 1 and System 2 thinking are not exclusive modes of cognition that function independently [149]. These modes of thinking form a dynamic partnership with the goal of making better quality decisions — within the context of this dissertation, decision making is analysed in this light. In the remainder of this section, various judgement heuristics that form part of human decision making, as well as the cognitive biases that render heuristic judgement imperfect, are investigated.

3.2 Judgement heuristics

The human ability to make judgements — to perceive stimuli, filter out relevant information, assess that information according to some internal criterion, all the while maintaining applicability to the context under consideration — is certainly quite remarkable. Over the years many philosophers have pondered this ability, weighing its effectiveness and the moral foundations of its assessment. As these questions remain of interest, this section is devoted to an inquiry into the practical mechanisms of human judgement.

Toward that end, the following logical principle is noted: The nature of the environment within which a decision is made constitutes the determinants of the strategies that can be deemed appropriate. Accordingly, the uncertain nature of the real world results in most human decisions being based upon a number of beliefs with respect to the probabilistic likelihood of uncertain events [456]. Typical examples of such decisions include the outcome of an election, the future value of the South African Rand, or the guilt of a defendant in a court case. It is simply impossible to know the outcome of these kinds of real-world events with probabilistic certainty, a fact that begs several questions of importance, such as: *What are the origins of the beliefs that dictate individual decision making?* or *By which mechanisms do people assess the likelihood of uncertain events [456]?* These questions underlie the theme of this section: The manner in which people rely on a limited number of heuristic principles aimed at simplifying the complex task of predicting values and assessing likelihoods. In general, these heuristics are useful, but occasionally they lead to severe and systematic errors [68].

Various kinds of heuristics are considered in this chapter. Consider a mortgage banker, for example, who might follow the heuristic of spending only 35% of income on property, or the poker player who might follow the heuristic of never playing for an inside straight. Although an understanding of these specific heuristics is important to the aforementioned professionals, their applicability falls outside the scope of this chapter. This chapter is dedicated to a review of the more general cognitive heuristics that affect virtually everyone. These general heuristics are partitioned into three categories, namely: *Simple heuristics*, *complex heuristics* and *hyper-heuristics*. Simple heuristics refer to those heuristics that constitute a single rule of thumb used by individuals as a guiding principle in decision making. Complex heuristics, on the other hand, are those heuristics that utilise multiple rules of thumb (*i.e.* simple heuristics) and intricate cognitive processes during judgement actions and decision making. Finally, hyper-heuristics represent the class of heuristics that approach decision making by selecting, combining and adapting several simple and complex heuristics. These categories are not formally recognised in the literature, but are suggested by the author as a beneficial classification of heuristics within the context of this dissertation.

Finally, it should be intuitively apparent that there are many simple heuristics. There are, in fact, too many of these heuristics to document effectively. Inspired by the academic literature, the predominant focus of this section is thus dedicated to the working of complex and hyper-heuristics, although a select few simple heuristics are also discussed. More specifically, the following hyper-heuristics are considered: The *affect* heuristic and the *simulation* heuristic. Thereafter, the following complex heuristics are considered: The *availability* heuristic, the *representativeness* heuristic, the *anchoring-and-adjustment* heuristic, and the *recognition* heuristic. Finally, the following simple heuristics are considered in closing: The *similarity* heuristic and the *effort* heuristic.

3.2.1 The affect heuristic

Recognition of the affect heuristic [424] is probably one of the most important developments in the study of judgement heuristics during the past few decades. A strong, early proponent of the importance of affect in decision making was Zajonc [493], who argued that affective reactions to stimuli are often an individual's very first reaction, occurring automatically and guiding subsequent information processing and judgement. As used here, the word "affect" refers to the specific quality of "goodness" or "badness" experienced as a feeling state.

At present, there is compelling evidence in the literature confirming Zajonc's initial proposition that every stimulus elicits an affective evaluation that occurs before any higher level reasoning

takes place and frequently forms the basis of subsequent decision making (see the reviews by Bargh [45] or Zajonc [492, 493] in this respect). This phenomenon by which decision makers utilise their affective response as a basis for subsequent decision making captures the essence of the affect heuristic.

As is often the case with anything involving emotions, the affect heuristic does not exhibit the features of constituent elements of an exact science. There are a number of elements that determine the degree to which individual decision making is influenced by affect dynamic in nature. These elements include (but are not limited to): The characteristics of the decision maker, the characteristics of the decision making task, and the dynamic interaction between the decision maker and the task. In other words, individuals differ in the way they react affectively and the extent to which they are inclined to rely on affective feedback; while tasks differ with respect to their evaluability (*i.e.* the relative affective salience of information) [424].

3.2.2 The simulation heuristic

During the latter half of the 20th century, computerised simulation techniques have become one of the dominant methods of optimisation. Interestingly, however, within the realm of optimisation by simulation, the decision sciences have been rather late to the party — the mental models of human cognition have long since incorporated simulating capacities. Accordingly, the simulation heuristic describes the cognitive process of constructing mental models of reality in which hypothetical events take place [374]. These models may be constrained and controlled in several ways: The starting conditions can be left to equal their realistic default values, or can be modified to account for some special contingency. Furthermore, the model's outcomes may be left unspecified, or a target outcome may be set, in which case the model's objective would be to find a path to that state from the initial conditions [263].

Similar to a computerised simulation, a cognitive simulation does not necessarily have to produce a single story. In fact, the ease with which a model can produce different parallel outcomes is a critical criterion when judging the likelihood of multiple competing outcomes. More specifically, the ease with which the simulation of a system reaches a particular state can be utilised to judge the propensity of the real system producing that state [263]. Consider the following example which illustrates the working of the simulation heuristic.

During a study on the simulation heuristic, Kahneman [263] presented participants with the following scenario: "Mr. Crane and Mr. Tees were both scheduled to leave the airport at the same time, but on different flights. They travelled from a nearby town in the same limousine, were caught in a traffic jam and arrived at the airport thirty minutes after the scheduled departure time of their flights. On arrival, Mr. Crane was told that his flight had left on time, while Mr. Tees was told that his flight had been delayed, and had just left five minutes prior to his arrival. Who would be more upset? Mr. Crane or Mr. Tees."

It should come as no surprise that 96% of participants argued that Mr. Tees would be more upset. What is it that makes this stereotype so obvious? Note that the objective situation of the two gentlemen are identical — both have missed their flight. In every sense of the word, the difference between Mr. Tees and Mr. Crane is immaterial. Why then should Mr. Tees be more upset? Interestingly, Mr. Tees is expected to be more upset, because it would be easier for him to construct mental simulations in which he could have arrived five minutes earlier, than it would be for Mr. Crane to imagine how the thirty minute delay could have been avoided [263]. This example effectively illustrates the way individuals utilise the simulation heuristic to determine the likelihood of events, based on the ease with which mental models that lead to such events can be constructed.

3.2.3 The availability heuristic

The (complex) availability heuristic was established to capture one of the most widely shared assumptions in decision making and social judgement research: People estimate the frequency of an event, or the likelihood of its occurrence, by the ease with which instances of, or associations with, that event come to mind [455]. For example, a product manager may base his or her assessment of the probability that a new product will capture the imagination of the market on his or her recollection of the performance of similar products in the past [51]. This strategy would be effective in most cases, for the ease with which our minds recall instances of events is generally proportional to the actual frequency of that event [407].

The availability heuristic is, however, fallible — the availability of information in an individual's memory is influenced by multiple factors that are unrelated to the objective frequency of an event. For example, an event that evokes emotions is easily imagined and will be more easily available in memory than an event that is unemotional, bland, difficult to imagine, or vague [51]. Similar to the other heuristics discussed in this section, the fallibility of its nature renders the availability heuristic a tool that should be understood and employed with caution.

3.2.4 The representativeness heuristic

The literature describes the representativeness heuristic as being descriptive of an individual's cognitive processes when making judgements about an object or person by comparing them with representations, stored in memory, of other objects or people [183, 455]. Simply put, when making judgements about an individual, people tend to look for recognisable traits that correspond with previously formed stereotypes [456].

Consider an individual, Mr or Mrs X, who has been described as introverted, meek, meticulous and solemn. Given a set of occupational roles (*e.g.* salesman, farmer, pilot, librarian, physician), how would one evaluate the likelihood of Mr or Mrs X engaging in each of the respective occupations? When faced with these kinds of *linking* judgements, individuals typically employ the representativeness heuristic and proceed to measure the similarity between the given characteristics of Mr or Mrs X and the stereotype of each occupation, ordering the occupations by the degree to which Mr or Mrs X is an accurate representation of the respective stereotypes [51].

In some cases, like the situation described above, the representativeness heuristic can be a useful decision making tool, but this heuristic can also lead to serious errors in judgement. For instance, it took a long time for the medical society to accept the germ theory of disease [51]. Physicians found it difficult to accept the notion that viruses or bacteria could cause such devastating consequences as tuberculosis or measles. Instead, because they were relying on the representativeness heuristic to inform judgement, it was believed for centuries that disease was caused by malevolent agents, such as magic spells or evil spirits [51].

Interestingly, there seems to be a correlation in the literature between the representativeness heuristic and the study of transference within the field of social psychology. Transference refers to the phenomenon whereby assumptions and experiences learned during past relationships resurface in present-day relations [175]. In the socio-cognitive model of transference, this notion is defined in terms of the activation and use of a significant-other representation in an encounter with a new person [19]. The activated significant-other representation is typically utilised to make various assumptions about the encountered individual, not unlike the representativeness heuristic (see Chen *et al.* [105] for a discussion in this respect).

3.2.5 The anchoring and adjustment heuristic

Studies have shown that when asked, respondents typically do not know the exact year George Washington became the first president of America. They do know, however, that it had to have been after 1776, as the declaration of independence was signed in 1776. Thus, while calculating their answer, individuals tend to use the year 1776 (*i.e.* what they know with certainty) as a starting point. Respondents simply add a few years to their initial estimate until they reckon that they have reached a plausible answer [419].

In the face of uncertainty, as illustrated in the above example, individuals tend to make judgements by starting from an initial position that is then adjusted to yield a final answer. The initial position may be suggested by the formulation of the problem, the experience of the analyst or may be the result of a partial computation [456]. This tendency to utilise what is known as an anchor while exploring the unknown captures the essence of the anchoring and adjustment heuristic [143].

Like other heuristics that enable efficient decision making, the anchoring and adjustment heuristic can cause serious errors in judgement [143]. In one of the best-known anchoring studies [456], research participants were given a random number between zero and one hundred. The participants were then asked to indicate whether the percentage of African nations that were members of the United Nations were higher or lower then the respective value they were given. Results clearly illustrated that the final estimates were influenced by the random number provided initially [437]. This finding, that an absolute numerical judgement may be assimilated toward the standard of a preceding comparative judgement, has been replicated multiple times in the literature (see, for example, Cervone and Peake [95], Jacowitz and Kahneman [236], Northcraft and Neale [351], and Plous [362]).

3.2.6 The recognition heuristic

Every day people wander through streams of sights, sounds, odours, tastes, and tactile impressions, some of which are novel while others have been experienced previously. The cognitive mechanism for distinguishing between novel and previously experienced stimuli seems to be rather specialised and robust [120, 405]. The recognition heuristic utilises this intricate mechanism in order to determine the value of multiple, competing stimuli. Furthermore, the following rule of thumb is utilised in this sifting process: If one of two objects is recognised, then the conclusion is drawn that the recognised object has a higher value with respect to the relevant criterion [191]. Consider the following example of this basic assumption.

In Sir Arthur Conan Doyle's *The Final Problem*, Sherlock Holmes finally faces his arch enemy, Professor Moriarty. While describing Moriarty to Watson, the following conversation takes place: "You have probably never heard of Professor Moriarty, have you Watson?" "Never." "Aye, there's the genius and the wonder of the thing! The man pervades London, and no one has heard of him. That's what puts him on a pinnacle in the records of crime" [354]. Sherlock Holmes is clearly implying that extraordinary things can usually not avoid being known and thus, being recognised. The fact that Watson had not heard of Professor Moriarty would have been much less surprising if he was insignificant. In other words, it is assumed that recognition is often proportional to importance with respect to a specific criterion [354].

The recognition heuristic needs to be distinguished from the availability heuristic, as described in §3.2.3 [455]. The availability heuristic is based on recall, not recognition. People are able

to recognise far more items than they can recall [354]. Furthermore, availability is a graded distinction between items in memory that is measured by the ease with which instances of, and associations with, an event come to mind [455]. The recognition heuristic, on the other hand, does not address comparisons between items in memory, but rather the difference between items in and out of memory [190].

3.2.7 The similarity heuristic

The *law of similarity*, as a general principle, captured the attention of modern psychologists during a study on the response of educated Americans (*i.e.* undergraduates) to disgusting stimuli [391]. As described by Frazer [171] and Mauss [226], the law of similarity can be defined by the following statement: *Like causes like*, which essentially holds that causes resemble their effects. This statement forms the basic assumption of the (simple) similarity heuristic that individuals utilise in judgement and decision making.

An example of the "like causes like" principle can be found in modern medicine. The modern medical predilection is to assume that the treatment of a disease will have some surface relation to the presumed cause of the disease. Thus, in order to discover a cure for a disease, pharmaceutical scientists initially focus on the cause of the disease in the hope that the cause will illuminate or suggest the cure [188].

Interestingly, there seems to be a strong overlap between the similarity heuristic and the representativeness heuristic [456], as described in §3.2.4. During application of the representativeness heuristic, an entity or event is assigned to a category based on the similarity of its principal features with other members of that category. If the category in question is causes and their effects (*e.g.* the process of identifying the damage to a vehicle by analysing the symptoms), the similarity heuristic seems to partner with the representativeness heuristic and work in tandem.

3.2.8 The effort heuristic

The notion that effort influences judgement is an old and important idea in social psychology [275]. Some of the earliest work in dissonance [157] and self-perception [58] demonstrated that the more effort individuals invest, whether in the form of time, money or physical exertion, the more positively they evaluate the product of that effort [275]. In other words, individuals utilise effort as a heuristic to determine quality.

A common phenomenon in an education system is that students often complain about poor grades. Occasionally, such protests stem from errors in the grading system or strategic attempts by students to "work" the system, but normally, the grades are correct, and the protest sincere [275]. As a matter of fact, this protest is a result of the effort heuristic — the student's inherent assumption that because he or she had invested a great deal of effort, the outcome of the effort should have been great as well.

As the previous example illustrated, the quality of an item can be difficult to determine. For example, the quality of a managerial decision, the theoretical contribution of an academic manuscript or the monetary value of a painting, can all be difficult qualities to assess. Consequently, individuals utilise effort as an indicator of quality [275]. All else equal, paintings that receive the lion's share of an artist's attention tend to be of a superior quality when compared with those that receive less attention. Note, however, that the relationship between effort and

quality is inconsistent. As any artist knows, a moment of inspiration can occasionally rival what would otherwise take hours to produce.

3.3 Cognitive biases

The human mind carries spectacular power in its modest one and a half kilogram of mass. With little effort, people are able to accomplish sophisticated tasks, such as recognising faces or catching a ball. Most people are, however, largely unaware of how their minds accomplish these complex tasks. This lack of insight into the dynamics of our minds might seem insignificant, but such ignorance has the potential to result in profound consequences [51]. Without an understanding of the dynamics that drive decision making, it is impossible to anticipate when cognitive processes, which normally serve us well, are leading us astray. The focus in the previous section was on a number of judgement heuristics employed in pursuit of understanding decision making. This section furthers that pursuit with a review of a number of *cognitive biases* that render these heuristics imperfect.

A *cognitive bias* refers to the systematic pattern of deviation from rationality in judgement and decision making, whereby assumptions about other people or situations are made in an illogical fashion [205]. There are many cognitive biases [51, 68, 205, 456]. In fact, there are too many to investigate exhaustively in detail. In the following section, only the most relevant cognitive biases are discussed. Furthermore, mimicking the general approach in the academic literature, these biases are reviewed with respect to the heuristics they influence. Table 3.1 contains a synthesis of all documented cognitive biases the author could uncover in the literature.

3.4 Biases emanating from the affect heuristic

There is no dearth of evidence to substantiate the claim that individuals perceive reality according to two distinct modes of thinking — System 1 and System 2 thinking (described in §3.1) [425]. While these modes of thinking function interdependently, they remain fundamentally unique. A defining characteristic of System 1 thinking is its affective foundation [424] and, more specifically, its employment of the affect heuristic. Section 3.2.1 was dedicated to the affect heuristic and contained a strong case in favour of its use, portraying it as the centrepiece of the System 1 mode of thinking. Like other heuristics, however, reliance on the affect heuristic can also lead one astray. Although not formally coined as judgment heuristics, several phenomena documented in the literature render this heuristic imperfect. These phenomena are discussed in the remainder of this section.

3.4.1 Manipulation of affect in our daily lives

Being vulnerable to affective influence is the doorway to many beautiful human qualities, such as empathy and altruism. The danger of affective manipulation, however, is that it often takes place subconsciously — rendering us slaves to its influence [425]. Consider, for example, the subtle influence of affect-laden tags, such as “new,” “organic,” and “improved,” often printed on commercial products, or that of the affectively pleasing stage-names entertainers tend to adopt. One wonders whether the careers of John Legend, Avicii and Bruno Mars would have been as successful if they were to perform under their real names — John Stephens, Tim Bergling and Peter Hernandez, respectively. Objectively, these affectively pleasing tags and names do not add any kind of value, yet they dynamically influence our judgements [425].

TABLE 3.1: *List of cognitive biases.*

Bias	Description	References
Failures of the affective system	The tendency to overvalue affectively pleasing stimuli.	[158, 295, 423, 425]
The optimistic bias	The phenomenon where people's predictions about future events tend to be optimistically biased.	[77, 188]
The ease of recall bias	The tendency to overestimate the frequency of an event because of the vividness, recency or familiarity of a memory associated with an event.	[51, 407, 456]
The retrievability bias	The tendency to overestimate the frequency of an event because of the structure of the search set.	[51, 457]
The illusory correlation bias	The tendency to overestimate the frequency of co-occurrence with respect to phenomena that are similar to predetermined natural associations.	[97, 455]
The insensitivity to prior probability of outcomes bias	The tendency to ignore base rate information.	[11, 51, 261, 329]
The insensitivity to sample size bias	The tendency to underestimate variation in small samples.	[43, 151, 260, 456]
The misconceptions of chance bias	The tendency to view chance as a self-correcting process in which a deviation in any direction induces a deviation in the opposite direction.	[51, 59, 260, 456]
The illusion of validity bias	The phenomenon whereby people make categorical predictions, but fail to question the validity of input information.	[261, 260, 456]
The insensitivity to predictive accuracy bias	The phenomenon whereby people make numerical predictions, but fail to question the validity of input information.	[261, 260, 456]
The misconceptions of regression bias	The belief that outcome should be maximally representative of input, which results in inaccurate future predictions.	[181, 209, 261, 456]
The anchoring bias	The tendency to rely too heavily on a single piece of information during decision making.	[144, 143, 336, 456]
The belief bias	The tendency to evaluate the logical strength of an argument by the believability of its conclusion.	[150, 271, 303, 342]
The confirmation bias	The tendency to ignore information that contradicts one's beliefs.	[272, 363, 344]
The conjunction fallacy	The tendency to assume that more specific conditions are more likely than general ones.	[363, 457]
The distinction bias	The tendency to view two options as more distinct when evaluated simultaneously than when evaluating them separately.	[225]

TABLE 3.1 (continued): *List of cognitive biases.*

Bias	Description	References
The framing effect	The phenomenon where people draw different conclusions from the same information depending on how it is presented.	[154, 286, 299]
The hindsight bias	The tendency to overestimate the predictability of an outcome after it has occurred.	[24, 110, 363]
The illusion of control bias	The tendency to overestimate one's influence over external events.	[140, 448]
The negativity bias	The psychological effect by which individuals have a greater recall of unpleasant memories than positive ones.	[89, 197, 392]
The omission bias	The tendency to judge harmful actions as less moral than equally harmful omissions.	[48, 381]
The outcome bias	The inclination to judge a decision according to its eventual outcome instead of by the quality of the decision making process.	[47, 202]
The overconfidence bias	The general tendency to have excessive confidence in one's personal solutions to problems.	[66, 139, 363]
The consistency bias	The tendency to incorrectly remember one's past thoughts, perspectives, values and behaviour as resembling one's current self.	[83]
Cryptomnesia	A form of misattribution where memory is mistaken for imagination.	[403]
The fading affect bias	The phenomenon where emotions associated with a negative memory fades quicker than emotions associated with a pleasant memory.	[380, 472, 473]
The peak-end rule bias	The tendency to evaluate an experience not by its sum, but according to the average of what it was like at its peak and when it ended.	[257, 321]
The picture superiority effect	The manner in which concepts that are learned by visual imagery are more readily and frequently recalled than concepts that are learned by reading.	[126, 308, 416, 482]
The availability cascade	A self-reinforcing phenomenon in which a collective belief gains plausibility through its repetition in public discourse.	[276, 364]

3.4.2 Failures of the affective system

The failures of the affective system are not exclusively the result of external manipulation. The affective features that become salient in a judgement or decision making process can be faulty as a result of a number of inherent, internal, uniquely affective biases [424]. One such bias seems to surface when the affective system is tasked with the responsibility of discriminating between minor and substantial changes in the environment. For example, the affective system seems wired to sensitize individuals to small changes in their environment, at the cost of making them less able to respond appropriately to larger changes in their environment (*e.g.* the difference between 0 and 1 deaths in some accident, compared with the difference between 570 and 670 deaths) [425]. Fetherstonhaugh *et al.* [158] referred to this insensitivity to the magnitude of change as "psychological numbing."

Similar to magnitude, time has an interesting effect on the affect heuristic. It seems that decision makers are unable to appropriately evaluate changes that take place over a long period of time or are remote in time [425]. Today's hunger, pain and anger are palpable, but the same sensations anticipated in the future receive little weight [295]. The irrationality of the decision to start smoking is a prime example of such a failure to weight future consequences appropriately [423]. During a survey, Slovic [423] asked participants the following question: "If you had to choose all over again, would you start smoking?" More than 80% of young smokers (ages 14–22) and about 85% of adult smokers answered "no," dramatically illustrating the failure of the affective system to account for the impact of change that takes place over a long period of time.

3.5 A bias emanating from the simulation heuristic

Muhammad Ali captured the grandeur of our ability to imagine when he said, "The man who has no imagination has no wings." In addition to its inspiring potential, the human imagination is often employed in rather mundane tasks. One such task is the assessment of the frequency of an event whose instances are not stored in memory, but have to be generated (*i.e.* simulated) according to a given rule [263]. In such instances, an individual would typically employ the simulation heuristic, as defined in §3.2.2, by generating several hypothetical outcomes and evaluating the likelihood of those outcomes by the ease with which they were generated [456]. Unfortunately, the ease with which outcomes are generated does not always reflect their true frequency. This mode of evaluation is susceptible to several biases, one of which is the *optimistic* bias.

The optimistic bias is at the heart of the most robust findings in the psychology of predictive judgement, which is that people's predictions about future events tend to be optimistically biased [188]. Although it is a wonderful thing to be optimistic, it has been found by a number of metrics and across a variety of domains that people assign higher probabilities to the attainment of desirable outcomes than either logical analysis or objective criteria would warrant. Empirical studies of the planning fallacy have, for example, shown that people expect to complete projects in less time than it actually ends up taking [77]. This inherent sense of optimism which deters a more realistic simulation of future events can result in unpleasant consequences.

Surveys concerning the risk of falling victim to auto-mobile accidents [386], crime [479], and serious disease [203] have found that people believe their risk is less than average. While this belief may result in exaggerated feelings of safety and security, it also inhibits precaution. Evidently, too many people recognise the danger of certain behavioural patterns only after they fall victim to these patterns.

3.6 Biases emanating from the availability heuristic

In general, availability is a useful tool for determining the frequency or probability of an event. The availability of an event in memory is, however, influenced by a number of dynamics other than the frequency and probability of that event [456]. Consequently, relying on the availability heuristic leads to predictable errors in judgement (*i.e.* biases), some of which are reviewed in the remainder of this section. These biases include the *ease of recall* bias, the *retrievability* bias, and the *illusory correlation* bias.

3.6.1 The ease of recall bias

In pursuit of identifying the dynamics of the availability heuristic, Bazerman [51] presented participants with a list stating the five most common causes of death in the United States during the ten year period from 1990 to 2000. The following causes were listed: Tobacco, poor dietary habits and physical inactivity, motor vehicle accidents, firearms, and illicit drug use. The participants were asked to rank these causes of death from most common to least common.

In general, when tasked with such an assignment, an effective strategy would be to utilise the availability heuristic and attempt to retrieve instances (in memory) of the respective causes of death, consequently ranking the causes proportionally to the ease with which instances of the respective causes were retrieved [455]. In this instance, however, this strategy caused participants to underestimate the likelihood of death due to tobacco and poor diet, while overestimating the number of deaths resulting from motor vehicle accidents, firearms and illicit drug use [51]. It may come as a surprise that the causes of death listed in the previous paragraph are, in fact, listed in order of frequency [328]. Why then, would participants conclude otherwise? This deviation in judgement is a result of the vividness that is often associated with deaths resulting from motor vehicle accidents, firearms or drugs. In the same way that seeing a house burn down will have a greater impact on a subjective estimate of the probability of such incidents than merely reading about a fire in a local paper, the increased vividness of deaths resulting from motor vehicle accidents, firearms and illicit drug use increased the availability of such instances in the memories of the participants. This increased availability is, however, not proportional to the objective likelihood of such events [51].

In addition to vividness, there are other factors that affect the retrievability of an event. For example, recent events are expected to be more easily available in memory than earlier events. Furthermore, the familiarity of the item being retrieved from memory dramatically influences its retrievability. In an elementary demonstration of this phenomenon, Tversky [456] exposed subjects to lists of well-known male or female celebrities. In some of the lists, the women were relatively more well-known than the men, and in others the men were relatively more well-known than the women. After being read out loud, participants were asked to judge whether a list contained more male or female celebrities. In every instance, participants incorrectly judged the class consisting of the better known celebrities to be the more numerous [456].

3.6.2 The retrievability bias

Suppose a word is sampled at random from a selection of the English literature. What is the probability that the word will start with the letter r, compared with r being its third letter? Individuals typically attempt to solve this problem by recalling words that begin with r (*e.g.* rabbit) and words that have r as its third letter (*e.g.* bar). They consequently judge the relative

frequency of the respective cases to be proportional to the ease with which the two types of words come to mind [455].

Note the potential for bias in the above strategy. People are better at retrieving words from memory using a word's initial letter than a word's third letter [51]. This is simply a result of the way in which the human brain structures information in memory. Tversky and Kahneman [457] demonstrated this retrievability bias beautifully, by showing that individuals estimated the frequency of seven-letter words ending with the more memorable suffix *ing* to be greater than the frequency of seven-letter words with the letter n in the sixth position. This is clearly incorrect as all seven-letter words ending with *ing*, have the letter n in the sixth position. This example illustrates the subtle way in which the structure of information in memory, and the structure of the world itself, interact to weaken one's search strategies, establishing the origin of the retrievability bias.

Note that the retrievability bias is fundamentally different from the ease of recall bias discussed in the previous section. While the ease of recall is determined by the character of that which is being recalled, the ease of retrieval is determined instead by the interaction between the structure of the brain and the structure of things searched for. In other words, the retrievability bias does not regard the emotional intensity of a memory, for example, but only considers the structure of that memory in the brain and how that structure influences the ease of its retrieval, deeming the documentation of both biases appropriate.

3.6.3 The illusory correlation bias

Chapman and Chapman [97] chronicled an interesting bias in the judgement of the frequency with which two events co-occur. In a particular study, Chapman and Chapman presented inexperienced participants with the clinical diagnoses and test material of several hypothetical patients. After a given time during which participants could study the respective documents, they were asked to estimate the frequency with which each diagnosis (*e.g.* suspiciousness or paranoia) had been accompanied by certain symptoms (*e.g.* peculiarities in the expression of the eyes). The results indicated that subjects overestimated the frequency of co-occurrence with respect to phenomena that were similar to predetermined natural associations, such as suspiciousness and peculiar eyes. This effect was labelled illusory correlation. The illusory correlation effect proved exceptionally resistant to contradictory data and prevented participants from identifying relationships that were actually present [97].

The availability heuristic provides a natural explanation for the illusory correlation effect. The boundaries between the specific instance being investigated and the more generally available memory are blurred. Judgement of how frequently a diagnosis is accompanied by certain symptoms in a specific instance is biased by the more general associations that are predetermined (more available) in memory. According to this view, the illusory correlation between suspiciousness and peculiar eyes, for example, is a result of the fact that suspiciousness is readily associated with an individual's eyes [455].

Life-long experience teaches one that the associative connections between events are strengthened when they co-occur frequently. As a result, individuals have developed fitting strategies to determine the co-occurrence frequency of events by the ease with which relevant mental operations of association can be performed. As the preceding examples demonstrate, however, this valuable decision making procedure is vulnerable to systematic errors.

3.7 Biases emanating from the representativeness heuristic

First described by Tversky and Kahneman in the 1970s [455], the representativeness heuristic is a decision making strategy that employs the use of past experiences to guide decision making. More specifically, the representativeness heuristic is employed when individuals make judgements about an object or person, by comparing them with representations, stored in memory, of other objects or people (see §3.2.4 for a more comprehensive description) [183]. This approach to judgement and decision making may lead to serious errors — several factors that should be taken in account when evaluating likelihood plays no role in judgements of similarity. Notably, the following biases are considered: The *insensitivity to prior probability of outcomes* bias, the *insensitivity to sample size* bias, the *misconceptions of chance* bias, the *insensitivity to predictive accuracy* bias, the *illusion of validity* bias, and the *misconceptions of regression* bias.

3.7.1 The insensitivity to prior probability of outcomes bias

One of the elements that has a major effect on probability, but no effect on representativeness is the prior probability of an event. Consider once again the example of §3.2.4 in which an analyst was asked to identify the day job of a Mr or Mrs X from a set of occupational roles (*e.g.* farmer, librarian, pilot, salesman or physician) based on a brief personality portrait. The fact that there are many more farmers than librarians in the general population, for example, should have entered into any sensible estimate of whether Mr or Mrs X is a farmer or librarian [11]. Unfortunately, however, considerations of the similarity between Mr or Mrs X and typical stereotypes pay no attention to prior probability, rendering the typical analysis incomplete. Ignoring prior probability in this way may result in many unfortunate implications [51]. Prospective entrepreneurs, for example, typically have far too much confidence in their future success and far too little appreciation for the prior probability of business failure. Many of these entrepreneurs end up loosing their life savings as a result [329].

In an experiment that related to the prevalence of the prior probability bias, Kahneman and Tversky [261] presented participants with the following excerpt: “Dick is a 30 year old man. He is married with no children. A man of impressive ability and high motivation, he promises to be quite successful in his field. He is well liked by his colleagues.” The participants were furthermore told that the group from which this description had been sampled consisted of seventy engineers and thirty lawyers. The description had been designed to convey no information relevant to the follow-up question, which was whether Dick was an engineer or a lawyer. Consequently, participants were expected to judge the probability of Dick being an engineer to be proportional to the distribution of engineers in the group. The participants, however, judged the probability of Dick being an engineer to be 0.5 regardless of the stated proportion of engineers in the group being 0.7.

Evidently, people are surprisingly insensitive to the prior probability of an outcome. And although the representativeness heuristic is a powerful decision making tool, its weaknesses should be kept in mind.

3.7.2 The insensitivity to sample size bias

Errors and biases in judgement under uncertainty are a major source of data for mapping the boundaries of a person’s statistical intuition. In a series of ingenious experiments, Kahneman and Tversky [260] elucidated a section of this boundary by demonstrating people’s insensitivity

to the effects of sample size on probability. Subjects were presented with the following example: "A certain town is served by two hospitals. In the larger hospital about 45 babies are born each day, and in the smaller hospital about 15 babies are born each day. As you know, about 50% of all babies are boys. The exact percentage of baby boys, however, varies from day to day. Sometimes it may be lower than 50%, sometimes higher. For a period of one year, each hospital recorded the days on which more than 60% of the babies were boys." Which hospital do you think recorded more such days?

Most of the participants (56%) judged the number of recorded days to be the same for the small and large hospitals. In contrast, sample theory holds that the expected number of days on which more than 60% of babies were boys, should be much larger for the small hospital. Essentially, a large sample is less likely to deviate from the mean (*i.e.* 50%) [43]. This fundamental notion in statistics is evidently not part of human intuition. Furthermore, this insensitivity to sample size seems to persist in judgements of posterior probability (*i.e.* the probability that a sample has been drawn from a certain population) [260, 456].

Interestingly, the academic literature is not coherent with respect to the insensitivity to sample size bias. Several researchers, including, Evans and Dusoir [151] and Bar-Hillel [43], repeated the above experiments only to arrive at results that were not consistent with that of Kahneman and Tversky [260]. Evans and Dusoir claimed that the tasks employed in Kahneman and Tversky's experiments were too complex to support an adequate test of the role of sample size. Bar-Hillel, on the other hand, did not provide specific critique, but also failed to replicate Kahneman and Tversky's original results [43]. In spite of these conflicting findings, the greater portion of the literature does seem to support the original findings of Kahneman and Tversky.

3.7.3 The misconceptions of chance bias

> *"Once we realize that imperfect understanding is the human condition, there is no longer shame in being wrong, only in failing to correct our mistakes." — George Soros*

Our imperfect understanding of the world is, at least in part, the product of inherent inconsistencies in the strategies the human mind develops in order to cope with the unpredictable nature of the world. A case in point is the misconceptions of chance that are generally present in the intuitive analysis of statistics [59]. Evidently, people incorrectly expect that a sequence of events generated by a random process will represent the essential characteristics of that process, even when the generated sequence is short [456].

The following question illustrates this misconception: A married couple has had three girls together. They are now expecting a fourth child, what is the chance of having another girl? Most individuals have a strong intuitive sense that the probability of having four girls is unlikely. Therefore, they posit that the probability of having a fourth girl should be less than 50% [51]. The problem with this reasoning is that the gender determination of every newborn baby is an independent chance event; the sperm that determines whether the baby will be a boy or a girl has no memory of prior impregnation. This example parallels research by Kahneman and Tversky [260] who elucidated the fact that individuals expect a sequence of random events to look random.

Consider the toss of a fair coin, which can either land *heads* (H) or *tails* (T). People judge the sequence HTTHTH to be more likely than the sequence TTTHHH, which does not appear random. Furthermore, people expect the former sequence to be more likely than the sequence TTTTHT, which does not seem to represent the fairness of the coin.

Despite the simplicity of these examples, they convincingly illustrate the fact that people expect the essential characteristics of a process to be represented, not only globally, but also locally in each of its parts. In contrast to this belief in representativeness, a locally representative sequence deviates systematically from chance expectation [456]. In general, chance is commonly seen as a self-correcting process in which a deviation in any direction induces a deviation in the opposite direction in order to restore the equilibrium of the system. In reality, however, deviations are not corrected — as a chance process unfolds, their effect is simply diluted [456].

3.7.4 The illusion of validity & insensitivity to predictive accuracy bias

In their article, *Judgement under uncertainty: Heuristics and biases* [456], Kahneman and Tversky proposed a number of cognitive biases, including the illusion of validity and the insensitivity to predictive accuracy. After a comprehensive investigation, it has become apparent that these two biases are fundamentally similar. Both biases become relevant in the context of intuitive predictions (*i.e.* when the representativeness heuristic is employed to order outcomes according to the degree to which they represent the essential features of some input evidence) [261]. Both biases are rooted in the way people place a great deal of confidence in their predictions, despite the fact that the information upon which their predictions are based might be unreliable or outdated. Evidently, the only difference between the illusion of validity and the insensitivity to predictive accuracy is their context of application.

Kahneman and Tversky classified intuitive predictions into two discrete categories, namely *categorical* prediction and *numerical* prediction [260]. In the categorical case, the prediction is given in nominal form (*e.g.* an individual's future occupation, or the winner of an election). In the numerical case, the prediction is given in numerical form (*e.g.* a student's grade point average, or the future value of a particular stock). These two categories define the dividing line between the illusion of validity and the insensitivity to predictive accuracy — the former coming into play during categorical prediction and the latter manifesting itself during numerical prediction.

During categorical prediction, people typically place a great deal of confidence in their prediction that an individual belongs to a certain class when provided with a description of his or her personality which matches that of the class stereotype, despite the possible inaccuracy of information. For example, people place a great deal of confidence in their prediction that an individual is a librarian, given that a description of his or her personality matches that of a stereotypical librarian, despite the fact that the description may be unreliable or outdated. This unjustified confidence in predictive accuracy in the context of categorical prediction is called the illusion of validity [456].

Numerical prediction elicits a similar effect [261], the only difference being that the prediction is in the form of a numerical value. Suppose a subject is given a description of a company, and is asked to predict its future profit. If the description of the company provided is very favourable, a very high profit will seem most representative of that description. If, however, the description is mediocre, the representative profit will also be mediocre. Thus, the prediction is dramatically influenced by the nature of the description. The presence of a good fit between the description and future value, seems to cause individuals to ignore salient elements that should be considered when estimating the future value of the company. The unwarranted confidence that accompanies a good fit between available evidence and possible outcome in the context of numerical prediction is called the insensitivity to predictive accuracy [456].

3.7.5 The misconceptions of regression bias

The notion of *regression to the mean*, as coined by Galton [181], is one of the oldest in modern statistics, yet it is still regarded as somewhat mysterious. The mystery lies not within its mathematical definition, but rather in the driving forces behind its occurrence. The phenomenon that Galton christened "regression" is connected to the occurrence and measurement of the change in a particular characteristic from one sample to the next. More specifically, reference is made to the manner in which data points regress toward the mean of a sample from one generation to the next [456]. Galton originally noted that tall fathers tended to have tall sons. While taller than average, however, the heights of the sons tended to be less extreme with respect to the mean (*i.e.* the average height of the male population) than those of their fathers [209]. Galton described this dynamic by stating that the sons regressed toward the mean.

In more recent times, Kahneman and Tversky [456] investigated this idea of regression from the perspective of judgement under uncertainty. They argued that the misconceptions of regression were twofold: First, people do not expect regression in many instances where it is bound to occur. Secondly, when people successfully recognise the occurrence of regression, they typically invent spurious explanations for it [261]. Interestingly, Kahneman and Tversky were of the opinion that the phenomenon of regression remained elusive because it was incompatible with the popular belief that the outcome should be maximally representative of input [456]. The failure to recognise the prevalence of regression may result in detrimental consequences, as illustrated by the following example.

During the training of military fighter pilots, the successful execution of a complex flight manoeuvre is typically followed by a deterioration during the next attempt. Similarly, a failed attempt is typically followed by an improvement. These dynamics are in perfect harmony with the notion of regression toward the mean and will occur even in the absence of feedback from the instructor, although the instructor might conclude differently. Since flight instructors typically praise a trainee after a good performance, and reprimand him or her after a poor performance, they are liable to reach the false conclusion that punishment is more effective than positive reinforcement in shaping behaviour [456].

3.8 A bias emanating from the anchoring heuristic

The nature of everyday life often requires people to estimate quantities that are subject to uncertainty [142]. Fortunately the human mind has developed a number of strategies for coping with tasks of this nature. One such strategy, described in §3.2.5, is the anchoring and adjustment heuristic. Kahneman and Tversky [456], the pioneers of this heuristic, noted that in addition to its ability to manage uncertainty in decision making, it also led to predictable deviations from what is generally considered good judgement [143]. One such deviation is the *anchoring* bias.

In a demonstration of the anchoring bias, Kahneman and Tversky [456] instructed participants to estimate a number of quantities in the form of a percentage (*e.g.* the percentage of African countries that are members of the United Nations). For each question, a starting value (*i.e.* an anchor) between 0 and 100 was established by spinning a wheel of fortune. Unknown to the subjects, the wheel of fortune was rigged to stop at either 10 or 65 percent. After spinning the wheel, participants had to judge whether the given starting value, as determined by the wheel of fortune, was higher or lower than their estimate. Finally, participants had to state their final estimate. It may seem counter-intuitive, but the initial value (*i.e.* either 10% or 65%) had a profound influence on subjects' subsequent estimates. For example, the median estimates for

the percentage of countries in Africa that were part of the U.N. were 25% and 45%, for the group of participants receiving 10% and 65% as initial value, respectively [456]. In general, the starting information, or anchor, seems to exert a drag on the subsequent adjustment process, resulting in final estimates that are too close to the original anchor.

Researchers have suggested a number of viable explanations for these anchoring effects. Testing an idea first advanced by Quattrone *et al.* [370], Epley and Gilovich [143] obtained evidence that people adjust insufficiently from an initial anchor because they stop adjusting once their estimate falls within a certain range of reasonable values [144]. Therefore, individual estimates tend to lie on the anchor's side of this reasonable range, while, on average, the true value lies closer to the middle of the range. Mussweiler and Strack [336] provided another explanation for the anchoring bias by showing that the presence of an anchor elicits information from memory that is consistent with the anchor, resulting in a similar drag effect.

3.9 Chapter summary

Uncertainty has proven to be one of the few things one can be certain to expect when making the decisions that everyday life demands. This chapter was devoted to reviewing the critical nature of the need to understand and be aware of the cognitive dynamics that influence daily decision making. Ignorance of these dynamics is considered to be one of the root causes of poor decision making.

The stage was set for a deeper investigation into the dynamics underlying decision making under conditions of uncertainty by reviewing the notions of intuitive and analytical thinking in §3.1. Subsequently, §3.2 was devoted to a description of the heuristics that are commonly employed in judgements related to uncertain events. These heuristics were classified into three categories, namely simple heuristics, complex heuristics and hyper-heuristics. Finally, several cognitive biases that render judgement heuristics imperfect were described in §3.3–§3.8.

CHAPTER 4

Group behaviour

Contents

"Our ability to reach unity in diversity will be the beauty and the test of our civilization"
— *Mahatma Gandhi*

Unity within diversity is truly a virtue worth pursuing. This chapter is devoted to a review of topics in the academic literature that are related to the social dynamics of human decision making and deemed influential in pre-empting such unity. These topics are introduced in §4.1 in the form of a discussion about the paradoxical existence of group behaviour, and this is followed in §4.2 by a review of the dynamics internal to each individual member who forms part of a group. Thereafter, the dynamics that exist between these individual group members are discussed in §4.3. Finally, the phenomena that dictate intergroup interaction are discussed in §4.4, before the chapter closes with a brief summary in §4.5.

4.1 The curious case of group behaviour

In the context of Western culture, consistency is a desirable characteristic. People who are perceived as behaving consistently across a variety of situations are positively affirmed, while inconsistent individuals are often labelled "two-faced" or "insincere" [439]. As a consequence, Western theories of social behaviour and personality, such as theories of attribution [267], attitude change [31, 157], partner selection [442], and individual traits [322], to name but a few, assume the importance of consistency between two domains, one related to an individual's self-concept and the other to his or her behaviour. Given the centrality of consistency in theories of social behaviour and personality, it is curious that researchers of *group dynamics* paint a very different picture concerning the behaviour of individuals under the influence of social stimuli. In essence, much research on the dynamics that dominate behaviour in the context of a group suggest that individuals are relatively pliable when it comes to fitting in and conforming to the

expectations of others[1] [223]. This posits an interesting question: Given the assumed importance of consistency in behaviour in theories of social behaviour and personality, why are people willing to adapt their behaviour in the context of a group?

At the heart of this contradiction rests a conflict between two fundamental human motivations: The need to be a consistent, differentiated individual and the need to experience group belonging [223]. On the one hand, there is an abundance of evidence supporting the notion that a sense of belonging within the context of a group is a powerful source of motivation (see Baumeister *et al.* [50] for a thorough review of this notion) and that collective identities can be fundamental to a person's self-concept. On the other hand, there is equal evidence of a countervailing need to differentiate oneself from others in a manner that is consistent with a person's self-concept, a process that is considered to be essential in defining the self [466]. Interestingly, the tension within this paradox can be traced back to the building blocks of an individual's self-concept — the individual, relational, and collective self [105].

The areas of study that have contributed most to the discussion on the relationship between the individual, relational and collective selves are social identity theory [218, 219, 453] and its extension, self categorisation theory [454]. Advocates of these theories (collectively called the social identity perspective [223]) argue that an individual's sense of self is informed, at least in part, by the groups to which he or she belongs [105]. Consequently, an individual's self-concept consists not only of distinct memories, attitudes and behaviours that distinguish themselves from other individuals, but also includes collective memories, attitudes and behaviours received from the group to which he or she belongs [72, 408]. Furthermore, self-categorisation theory argues that the more individuals identify with a salient group identity, the more likely they are to view themselves and other group members as interchangeable exemplars of the group's personification [220]. This basic phenomenon is thought to underpin group influence and conformity. In summary, the literature seems to suggest that people have a fundamental need to belong to social groups, while at the same time having a fundamental need to differentiate themselves in a manner consistent with their self-concept. It is within the balance of these fundamental needs that the phenomena of group behaviour take place.

4.2 The individual in the group

Every group of people, among the billions of groups that exist at this moment in time, is a unique configuration of individuals, relationships and processes [170]. A study group at a college library, for example, will differ in a hundred nuanced ways from a rock band performing a well-loved song, a team on a soccer field, or a board of executives electing a new chief executive officer. These groups, however, all possess common properties and dynamics despite their distinguishing characteristics. It follows that when researchers study group behaviour, they are required to look beyond the unique qualities of the group to the qualities that seem to be consistent across most groups. One strategy in pursuit of finding these more general qualities is to focus on the individual in the group — all groups fundamentally consist of several individual members. In order to understand the dynamics of a group, it makes sense to investigate the dynamics that each of its members bring to the table. Notably, it is the interaction between the dynamics of every individual group-member with every other, that produces the predominant dynamics of a group [223].

[1] Asch [32], for example, found that people were willing to express judgements that were clearly wrong, if it meant social acceptance. Similarly, in a study on small-group processes, Deutch and Gerard [130] found that people would often keep their dissenting thoughts private for fear of social sanction, despite potentially disastrous consequences [241]. Another case in point is researchers of social identity who posit that when a relevant group identity is salient, people adapt to the norms and values of the group [444, 452].

A common approach when studying the dynamics of the individual[2], related to the personality and social psychology literature, is to base investigation on the premise that an individual's self-concept consists of three fundamental self-representations, namely the *individual self*, the *relational self* and the *collective self* [105, 106, 402, 408]. These three selves co-exist to form a *universal self*, such that a person can alternate between perceiving themselves as a distinct individual, a relational partner, or a collective group member [409]. In addition, each self is important and meaningful to human experience [61]. These self-representations, forming the pillars that constitute a person's self-construal, are considered in this section.

4.2.1 The individual self

In general, the individual self refers to aspects of the self that make a person unique and separate from others[3] (*e.g.* world-view, interests, values, traits, experiences, attitudes and behaviours) [411]. It is salient then, that such traits will influence the real-time behaviour of an individual in the context of a group. In fact, the difficulty of collective decision making often stems from the nuanced uniqueness of each group member's individual self [103]. A simple case in point is the way in which one person's "bad" is often another's "good," or one person's "freedom-fighting" another's "terrorism." These differences in perspective can be explained by the uniquely subjective nature of the individual self, and more specifically, the subjective nature of the criteria according to which such judgements of good or bad are made (which is embedded in the individual self). These criteria are the product of both an individual's genetic inheritance (*i.e.* the kind of person one is innately) and, more significantly, an individual's previous experiences of the world [103]. Consequently, to the extent that our genetic make-up and experiences are unique, our individual self will be uniquely biased accordingly [361, 402].

4.2.2 The relational self

The assertion that the relational conceptions of the self are central to understanding human behaviour is almost as old as the discipline of psychology itself [238]. While the individual self illustrates the nature of the self in isolation from others, it is necessary to further define the self in a way that involves a connection to, rather than a separation from, others. This is where the relational self (and the collective self, to be considered subsequently) becomes relevant. In order to define the relational self, it is necessary to draw on the notion of a *working self-concept*, which refers to the self-knowledge that becomes available to an individual during specific social interactions. For example, a person's self-knowledge in relation to his or her younger sibling may include idiographic attributes such as *fun-loving*, or *jokester*, as well as normative roles such as *authority figure*. An individual's self-knowledge that becomes relevant in this manner proceeds to influence his or her behaviour in both subtle and dramatic ways [304]. More specifically, a person can be expected to behave differently in dissimilar contexts, depending on the self-knowledge that becomes salient — an individual's entire array of self-knowledge is certainly not accessible all at once [305].

Utilising this notion of a working self-concept, notice that if a certain significant other were to behave in a way that caused an individual to feel loved, across a variety of interactions and over a significant period of time (especially in dyadic relationships), the characteristic "loved"

[2] Although there are alternative conceptions of the self in the academic literature, the stated conception is reviewed as it is most prevalent.

[3] In the context of this dissertation, the individual self represents the unique perspective from which, and the unique way in which, an individual approaches life.

would likely become part of that person's working self-concept with respect to the significant other [105]. Over time, the significant other would thus no longer have to behave in a way that caused that person to feel loved; they would simply feel loved as a result of the real or perceived presence of the significant other — which would activate the individual's working self-concept with respect to that significant other, which in this instance includes the experience of being loved [106]. The relational self can then be defined as the working self-concept that becomes available in the presence of a specific significant other. This includes one's emotional experience when relating to a significant other, the goals one pursues in relationship with the other, the self-regulatory strategies one adopts in interactions with the other, and the behavioural tendencies one exhibits in the presence of the significant other [105]. This conception of the relational self correctly illustrates that people have several relational selves.

An insertion here is necessary. To state that the self-knowledge which becomes available in a social interaction is simply the product of the social environment, would incorrectly suggest that a person's social behaviour can be likened to that of a chameleon. A more accurate way of understanding these phenomena would be to claim that a person has a relatively consistent *global relational self* which constitutes the working self-concept from which an individual approaches the majority of social interactions [409], but that this global relational self can be overridden or influenced if specific social stimuli activate a different, more context-specific, working self-concept (*i.e.* relational self). Consider the following hypothetical example: Johnny has a highly critical mother. Johnny's global relational self can be described by words such as considerate, competent, and feelings of confidence, while his relational self with respect to his critical mother is composed of conceptions of the self such as inferior, feelings of rejection, and the goal to please. It follows that if someone, or something within the social context of Johnny's daily life, were to remind him of his mother and successfully activate his relational self with respect to his mother, he would experience himself like he would in the presence of his mother. This kind of altered experience of the self can dynamically influence an individual's behaviour [105]. Evidently, the relational self holds considerable potential in terms of influencing the real-time dynamics of a group.

In summary, the relational self is self-knowledge that is linked in memory to knowledge about significant others; exists at several levels of specificity; is capable of being chronically or contextually activated; and comprises self-conceptions as well as a number of other self aspects that characterise an individual when relating to significant others (see Chen *et al* [106] for a thorough discussion). Note, however, that the academic literature dedicated to the relational self elucidates several perspectives that are dissimilar to the one described above. The most dominant of these is the *inclusion-of-other-in-the-self* (IOS) approach [30]. Unlike the current conception which views the relational self as self-knowledge that is distinguishable from, but linked in memory to, knowledge about significant others, the IOS approach posits that close relationships involve the incorporation of significant others (*i.e.* perspectives and attributes) into the self-concept [29]. According to this view, the closer a relationship, the more the relational partner would be incorporated into the self. This perspective departs from the aforementioned view by suggesting that including others in one's self-concept reflects the internalisation of significant others into the self-concept, rather than conceptions of who one is in relation to significant others.

4.2.3 The collective self

The psychologist Carl Jung wrote about the collective unconscious [255], a source of deep, common knowledge and understanding that every individual seems to draw from. Although

a discussion on such a collective unconscious is left to more experienced thinkers, the notion illustrates the manner in which humans have always been drawn to a collective perspective of life, both philosophically and psychologically. In perfect harmony with this perspective, the collective self refers to a conception of the self as an interchangeable exemplar of some social category, rather than as being based on the unique attributes that define an individual as separate [411]. This form of self definition constitutes those aspects of the self that are based on some form of membership to a social group. Unlike the relational self, this membership is not based on personal bonds to others, but is derived from the shared identification with a specific group [410], that is, those aspects of the self concept that differentiate *ingroup* from *outgroup* members. For example, an individual may hold a collective identity as an environmentalist — whenever this aspect of the collective self is activated (similar to the activation of the relational self), similarities with other environmentalists are emphasised (*e.g.* a sense of responsibility toward the environment), while some traits related to the individual self (*e.g.* being an objective, logical thinker) may momentarily move into the background [302].

In several experiments, the collective self has been linked to an individual's behaviour toward other groups of which that individual is not a member (*e.g.* prejudice and inter-group stereotyping) [121, 450]. The collective self's potential for influencing the real-time behaviour of an individual is thus evident.

4.3 Group dynamics

The tendency to join with others in the format of groups is perhaps the single most important characteristic of humankind. History testifies to the fact that the processes which occur when individuals unite to form groups tend to leave an ineradicable imprint on their members and on society in general [169]. Despite its apparent influence, the dynamics within groups (*i.e.* the actions, processes, and changes that occur within and between groups [170]) have puzzled sages, scholars and lay people for decades. Why, it has been asked, do humans so frequently elect to join with others in groups? Which factors give rise to the sense of cohesion within groups, and which factors result in the distrust of individuals outside a group from within? And how do the leaders of a group hold sway over its members? Today, many of these questions have been answered, while some still pose challenging mysteries. Nevertheless, the inquiry into such questions has provided a scientific basis for the field of *group dynamics*, which is the academic discipline devoted to the study of groups and group processes [94].

Before reviewing the more intriguing topic of group dynamics, an interjection is necessary here. What is meant by a *group*? Interestingly, theorists are not of one mind when it comes to the definition of the word *group*. Some stress the importance of mutual dependence; others highlight the key role of communication. It has been suggested that a shared purpose is what turns a mere aggregate of individuals into a group [170]. Despite the ambiguity of its definition, there are several qualities specific to groups about which theorists do agree. In no particular order, these qualities include: *Interaction*, *goals*, *interdependence*, *structure*, and *unity*. Short descriptions of these qualities are provided in Table 4.1. Within the context of this dissertation, a group will be considered to be a social unit consisting of a number of individuals who stand in definite status and role relationships with one another and work toward a common goal within a bounded context.

The fact that humankind stands in debt to the accomplishments of collective groups of individuals deserves repeating. Without the collective contribution of individuals, humankind would

TABLE 4.1: *The qualities of a group, adapted from Forsyth [170].*

Feature	Description	References
Interaction	Groups establish, organise, and maintain *relationship* and *task* interactions among members	[40, 41]
Goals	Groups typically have an instrumental purpose, for they facilitate the achievement of outcomes sought by their individual members	[309]
Interdependence	There is a mutual dependence among the members of a group, in that each member influences, and is influenced by each other member	[170]
Structure	Groups tend to be organised, with each of their members connected to other members in a pattern of relationships, roles, and norms	[170]
Unity	Groups are cohesive social arrangements of individuals	[86, 132]

surely never have landed on the moon, built the Burj Khalifa, or have been able to perform one of Mozart's magnificent symphonies. An appreciation of these milestones in history should, however, not overshadow the fact that history also paints a less pleasant picture of group performance. For example, group decision making resulted in the *Challenger* and *Columbia* space mission disasters [156, 193], propped up the apartheid regime in South Africa [292] and contributed to the Deep Water Horizon oil spill in the Mexican gulf [385]. Individuals who come together to form groups have the potential of reaching great heights, but to ignore their limitations would be to invite disaster. In the remainder of this section, several dynamics that undermine group decision making are reviewed, namely *shared information bias*, *group polarisation* and the well-known *groupthink* phenomenon. A number of these dynamics are discussed and elucidated in the context of Case Study 4.1.

4.3.1 The shared information bias

After his inauguration, President JF Kennedy held the power to simply declare war on Cuba (as described in Case Study 4.1), but instead of acting independently, he formed a committee to help him review an invasion plan. Historians cannot confirm why Kennedy did this, but the theoretical reasoning behind utilising group decision making in critical instances is clear — the increased diversity of knowledge and opinion within a group is believed to enrich intellectual resources and better represent the diversity that is present within a typical community [464]. Why then, did this plan fail so miserably? One of the phenomena theorists have put forward to answer this question has been coined the *shared information bias* [170]. Researchers have found that participants of group decision making often fail to discuss information that is unique to individual members [487] and, furthermore, neglect to consider information that is contrary to the dominant opinion [406]. As it happened, the Bay of Pigs committee spent a considerable amount of time discussing the incompetence of the Cuban forces and how US citizens would react to the invasion, spending much less time talking about the more important operations-level military strategy of the envisaged mission, or the political climate in Cuba [170]. Several members of the committee, in fact, held information that may have altered the eventual decision. For example, the *Central Intelligence Agency* (CIA) representatives knew that the morale of the invasion force was very low, but kept this vital information to themselves. Furthermore,

CASE STUDY 4.1 *The Bay of Pigs Invasion, adapted from [170].*

The Bay of Pigs Invasion	
The US presidential election of 1960 saw JF Kennedy pitted against then vice-president RM Nixon. As presidential campaigns go, Kennedy selected several major issues to stress in his campaign, one of these issues being the spread of communism in the world, specifically referring to Cuba and the government of Fidel Castro. After a successful campaign, Kennedy faced one of his first major decisions: What to do about Cuba? The CIA responded by suggesting that the US should launch a covert military operation to topple Castro's government. Their plan was to send a squad of well-trained troops to capture and defend the strip of land in the Bahía de Cochinos (*i.e.* the Bay of Pigs) on the southern coast of Cuba. From there, the troops would perform tactical raids in order to spark a civilian revolt. Kennedy proceeded	to present the CIA's plan to the executive committee of the National Security Council. This committee included several White House senior advisors and staff members, a number of cabinet members, and representatives of the CIA, and the Joint Chiefs of Staff (*i.e.* military branch leaders) — all highly skilled individuals with expert training in making critical policy and military decisions. After a thorough review, this group advised the president to give the Bay of Pigs Invasion a go-ahead. On the 17th of April 1961 the Bay of Pigs Invasion took place. It was a total disaster. The entire US attacking force was either killed or captured within a matter of days. Group expert, Irving Janis, described the decision to invade Cuba as one of the "worst fiascos ever perpetrated by a responsible government" [240].

President Kennedy had information from several sources that would have forced the group to reappraise its decision, but he too kept this vital information to himself [274].

Stasser and Titus [434] studied this phenomenon by providing several four-member groups with information listing the performance of three candidates with respect to sixteen criteria. According to a rational analysis of the information, Candidate A was the best option with eight positive, four neutral, and four negative qualities. Candidates B and C both had four positive, eight neutral, and four negative traits. Initially, the soon-to-be group members were given all the information about each candidate separately, resulting in 83% of individuals correctly favouring candidate A. After group formation, however, the groups did not fair so well as the distribution of positive and negative information was manipulated so that a hidden profile emerged. Stasser and Titus ensured that each member of every group only knew of two of the eight positive traits of Candidate A. Participant 1, for example, knew that Candidate A possessed positive qualities Q_1 and Q_2, Participant 2 knew that candidate A possessed positive qualities Q_3 and Q_4, and so on. All participants, however, were informed about candidate A's four negative traits. The collective information within the groups was therefore no different from the initial experiment. The participants nevertheless failed to share the relevant information and selected an inferior candidate 76% of the time [434].

It is interesting to note that the shared information bias is not the consequence of collectively not holding information critical to a decision. It is rather the result of a segregation of information in a manner that results in a critical part of what is known, being neglected. Consequently, simply exerting greater efforts with respect to sourcing information will fall short of mitigating the shared information bias. Real-time group discussion strategies will be required to ensure that the shared information bias does not inhibit efficient and effective decision making. Toward developing such a strategy, it may be beneficial to further investigate the relevant underlying mechanisms.

In an academic society in which the relation between phenomena that are uncovered at different times by different researchers sometimes goes unnoticed, it is interesting to notice the potential relation between the shared information bias and the so called *Abilene Paradox* [204]. The Abilene Paradox can be described as the phenomenon whereby a preconceived idea of others' preferences cause individuals to pretend to hold similar preferences, resulting in a group reaching a false consensus with respect to an appropriate action plan [204]. In his article titled *The Abilene Paradox: The management of agreement* [204], Jerry B Harvey, the father of this paradox, recalls a memory in which he, his wife and in-laws spent an evening touring a neighbouring town, despite the fact that all of them secretly wanted to spend the night at home. Prior to the verbal discussion on whether or not to make the trip, each individual incorrectly assumed that everyone else wanted to go and consequently pretended to want to go too. Comically, each individual employed this strategy and the group ended up doing what they all would rather not have done.

This example illustrates how the Abilene Paradox seems to involve a section of what is known (*i.e.* not wanting to go to Abilene) being withheld, not unlike the shared information bias. The theory surrounding the Abilene Paradox suggests that the reasoning behind its occurrence can be summarised as a *fear of separation*, or in other words, social isolation. As described in §4.1, people have a fundamental need to belong. Therefore, if an individual believes that sharing an opinion contrary to the norm will lead to social isolation, he or she may be persuaded to remain silent. Thus, the one possible explanation for the shared information bias is the perceived belief that the majority of the group holds a position contrary to the implication of held information.

4.3.2 Group polarisation

Why President Kennedy created a committee to help him review the envisaged invasion plan will always remain a question open to debate. With that inherent uncertainty in mind, it is worth noting that President Kennedy may have acted on the assumption that a group, when faced with a choice between a risky alternative (*i.e.* invade Cuba) and a less risky alternative (*i.e.* utilise diplomatic means to influence Cuban politics), would favour the moderate option. As history portrays, however, this was not the case. Researchers have subsequently determined that groups do not necessarily nourish restraint; rather, they *polarise* — often making more extreme decisions than individuals [170].

The years leading up to the coining of the phrase *group polarisation*, was characterised by several confusing findings. Some researchers concluded that groups were more inclined to conservative decision options, while others found just the opposite [436, 474]. In order to measure this shift in a group's attitude toward decision options, researchers frequently employed the *Choice-Dilemmas Questionnaire* [170]. This self-report measure of willingness to make risky decisions can be broken down into the following parts. Initially, participants are presented with an extract such as the following:

> *"Mr Anderson is a qualified electrical engineer who has been working for a large electronics corporation since graduating from university five years ago. He is married with one child. His current employer has guaranteed him a lifelong job paying a modest, but adequate salary, including liberal pension benefits upon retirement. It is highly unlikely that his salary will increase substantially before retirement. During a local convention, a small, start-up company with a highly uncertain future offered Mr Anderson a job. This job would provide Mr Anderson with a better salary and the possibility of shared ownership if the company was to survive [170]."*

After reading the above excerpt, participants are asked to imagine themselves being Mr Anderson's advisor. Provided with the list of possible alternatives in Table 4.2, participants have to select the lowest probability of the start-up succeeding for which they would deem the move worthwhile.

TABLE 4.2: *The probability of success.*

Probability	Select
The chances are 1 in 10 that the start-up will succeed	
The chances are 3 in 10 that the start-up will succeed	
The chances are 5 in 10 that the start-up will succeed	
The chances are 7 in 10 that the start-up will succeed	
The chances are 9 in 10 that the start-up will succeed	
Mr Anderson should not take the new job regardless of the probability of success	

Several researchers employed this case study, or one similar in structure, with interesting results. In a manner that seems almost random at first, some researchers concluded that groups typically make riskier decisions, which they have dubbed the *risky-shift phenomenon* [368], while others found that groups consistently advocated a less risky course of action, dubbing this the *cautious-shift phenomenon* [474]. Intrigued by these inconsistent findings, subsequent researchers developed additional choice-dilemmas in order to further investigate these dynamics. They too, occasionally found evidence of both the cautious-shift and risky-shift phenomenon [133]. Eventually, researchers concluded that a shift toward either a risky or cautious decision attitude was possible, coining these collective phenomena *group polarisation* [133]. Curiously, the direction of the shift in attitude to either risky or cautious was determined to be a function of the initial average preference direction of the group [339]. Figure 4.1 is certainly an oversimplification of this process, but is appropriate for the purpose of illustrating the manner in which the direction of polarisation depends upon the initial state of attitudes in a group.

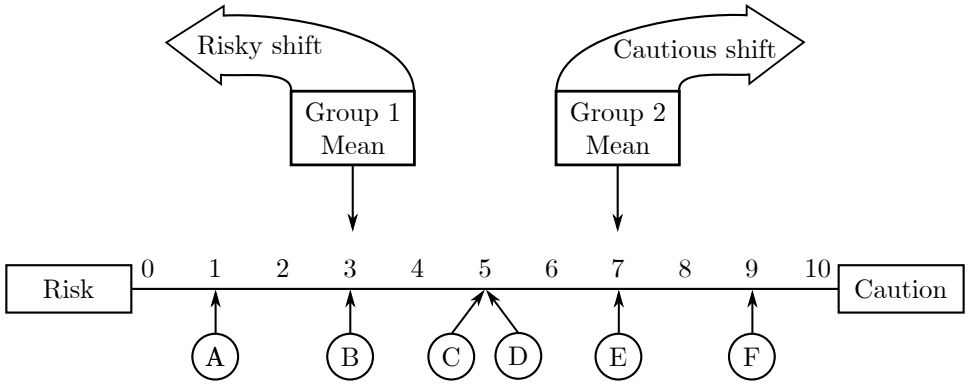


FIGURE 4.1: *A schematic representation of group polarisation, adapted from Forsyth [170].*

Imagine for a moment that President Kennedy had elected to split the gathered group of advisors into two independently functioning groups (Group 1 and Group 2) and had proceeded to discuss the invasion plan with each group separately. Suppose, furthermore, that Group 1 consisted of Individuals A, B, and C (whose attitudes toward invading Cuba were characterised by a 1, 3 and 5 on the sliding scale between Risk and Caution, respectively). Suppose also that Group 2 consisted of Individuals D, E and F (whose attitudes toward invading Cuba were characterised

by a 5, 7 and 9 on the sliding scale between Risk and Caution, respectively). This would have resulted in an average preference of 3 and 7 for Groups 1 and 2, respectively. Essentially, because the collective preference of Group 1 is on the risky side of the sliding scale, they would have been more inclined to polarise toward that end, while Group 2 would probably have polarised toward caution as a result of initially being more prone to caution collectively [338]. At this point, it is important to note that although the initial, collective decision attitude of a group is a good indication of the direction of polarisation, one should not assume that this trend will always hold true. It is, nevertheless, interesting to consider how important group composition may have been to the debate around invading Cuba.

The purpose of the discussion surrounding Figure 4.1 has been to illustrate that the direction of group polarisation is typically not random — it is rather a function of several underlying dynamics. One could say that the discussion demonstrated when groups will generally polarise on the side of risk or caution. At this point, however, the specific social dynamics which underlie group polarisation have yet to be elucidated. This leads to the next question of interest: What are the underlying social dynamics that cause polarisation to occur? This question has intrigued several researchers, of which, Liu and Latané [293] and Friedkin [176] were first to put forward the characterising phenomena of *persuasion*, *social comparison* and *social identity*.

Within the context of group discussions, persuasive argument theory holds that persuasive arguments which illicit an emotional response or an intellectual shift have a phenomenal potential to influence general opinion [176]. Contrary to this perspective, the social identity school of thought suggests that individuals are not persuaded by the content of arguments, but by an evident consensus of opinion [67]. In other words, if the members of a group come to believe that the prototypical group member holds a certain position, those who identify with the group (*i.e.* receive some form of social identity from being part of the group) will shift in the direction of that prototypical position [206].

Finally, social comparison theory holds that individuals spontaneously compare themselves with others, and if they identify a difference between their view and that of the group, they typically shift toward the group's view [401]. When people make decisions outside the context of others, they do not have the luxury of knowing how others would have responded to similar decision options. But in the context of a group, individuals can employ those around them as points of reference in order to evaluate their own position [189]. The presence of a measure against which to judge one's position may lead to the shift of position toward that measure [478].

The existence of the two schools of thought, persuasive argument theory and social identity theory, points toward a constituent of the academic society that proposes a complete divide between the influence of social identity and persuasive arguments on group discussions [170]. It is, however, suggested here, and supported by several researchers [467, 468], that there is a fundamental overlap between these influences. Groups typically generate more arguments that support the position endorsed by the majority of group members than arguments that oppose the generally accepted opinion [468] (similar to the shared information bias that was discussed in §4.3.1). As a result, there is an interesting circular influence during group discussion — it is within this circular influence that the overlap between persuasive argument theory and social identity theory becomes apparent. In general, individuals are more inclined to voice arguments that are in line with the social norms of the group. This one-sided tendency in argumentation makes the generally endorsed position seem even more dominant than it truly is. This, in turn, results in it becoming even more difficult to voice an alternative opinion as the risk of social isolation increases — a group essentially persuades itself [81, 80]. It follows that the interaction between social identity and persuasive arguments may lead to excessive group polarisation.

In conclusion to this section, consider the fact that group polarisation will not necessarily result in negative consequences. Depending on the issue at hand, group polarisation may, in fact, be of benefit if it results in a group exhibiting a greater conviction toward a shared goal in a context in which that goal would be beneficial. Consequently, one should be aware of group polarisation and its influence, but to extinguish its presence may not be wise. The way we as a people are influenced by the opinions of others, in the context of others, may be a unique way in which individuals are forced to consider alternative perspectives which may otherwise have been dismissed.

4.3.3 Groupthink

If it were appropriate, groupthink may have been deemed the celebrity of group phenomena. Several of the greatest failures in group decision making throughout history have been linked to the notion of groupthink, including France's defeat during WWII [10], the space shuttle Challenger disaster [146], the bombing of Pearl Harbor [239], and finally, the Bay of Pigs invasion [170]. President Kennedy's weary response after he and his advisers had blundered into the Bay of Pigs invasion was simply, "How could we have been so stupid?" This question intrigued the well-known social psychologist, Irving Janis, as he studied the situation in the context of several of the disasters listed above. Stupidity was certainly not to blame — the individuals who participated in the decision to proceed with the Bay of Pigs invasion represented one of the greatest assemblies of intellectual talent in the history of American politics [239].

After studying the groups involved in the aforementioned mishaps, Janis came to the conclusion that the presence of poor decision making was not the consequence of the separate dynamics of each individual present (*e.g.* temporary emotional states of elation, anger or fear that reduce an individual's mental efficiency), but rather a unique interplay between the members of a group [170]. Janis coined this interplay *groupthink*, which he defined as a manner of thinking in which people engage when they are deeply involved with others in a cohesive group, which typically results in members' desire for agreement overriding their motivation to realistically evaluate alternative courses of action [239].

At the time of its formulation, the notion of groupthink fundamentally challenged the prevalent perception in social psychology that a stronger sense of group cohesion and harmony resulted in improved decision effectiveness — the evidence that Janis uncovered suggested just the opposite [240]. In an article titled *Groupthink and France's defeat in the 1940 campaign*, Ahlstrom and Wang [10] explained that when a group is involved in high-stress decision making, strong group cohesion can contribute to an increased likelihood of decision error, as individuals may seek to downplay information that could potentially undermine the harmony experienced. The presence of harmony, however, can be likened to that of jelly on a peanut butter jelly sandwich — it's simply nice to have, but should not be allowed to dictate events.

Unlike jelly, however, the effects of groupthink are difficult to mitigate as it is not generally accessible through direct observation — the only credible way to establish its presence is by noticing the associated symptoms. Fortunately, Janis [240] held out a helping hand to future corporate professionals by defining eight symptoms associated with a group's beliefs and behaviour that could indicate whether groupthink was present. Notably, Janis employed the Bay of Pigs Invasion as one of the case studies to elicit these symptoms. The proposed symptoms can be related to one of three characteristics, namely an over-estimation of the group's capabilities, a closed-mindedness, and the presence of pressured conformity. The symptoms that relate to the foremost characteristic have been coined a belief in the *inherent morality* of a group, an *illusion of invulnerability*, and *collective rationalisations*. In the wake of these symptoms, a

group operates under the impression that its decisions will, by definition, be morally correct and virtually infallible [310]. Encapsulated by closed-mindedness, the symptoms that represent the second characteristic are a *stereotype of outsiders* and *self-censorship*. These symptoms result in an exclusion of information that does not coincide with the group's beliefs. Finally, an *illusion of unanimity*, a *direct pressure on dissenters*, and *self-appointed mind-guards* are the symptoms associated with the pressure toward conformity [240].

Robert Gates once said that no leader is well-served by everyone singing from the same sheet of music they think he's on [36]. This statement has proven to be true throughout the course of history. In order to ensure the effective functioning of our leading institutions, it is necessary to come to terms with and learn to mitigate the effects of groupthink and its associated mechanisms. Janis provided corporate professionals with an initial foothold in order to make the climb toward doing just that.

4.4 Intergroup dynamics

Social science first dipped its feet into the waters of intergroup relations thinking some time between the late eighteenth and early nineteenth centuries [280]. This era of research saw the birth of a field that would bring to bear a variety of methods and theories from the more general social sciences on a diverse set of intricate social problems [16, 360, 418]. Taken literally, intergroup relations refers to interactions *between* and activities *among* groups of individuals. Note that the choice of preposition is consequential. Whether researchers study groups only two at a time, or the interactions among more complex constellations, has a significant influence on what can be learnt. Furthermore, the study of intergroup relations attempts to explain a broader range of phenomena than just what takes place at the intersection between several groups. Studies range from investigations of how individuals think within the context of intergroup interactions, to how countries deal with one another in the realm of international relations [14]. Although the conflict that occurs between and among groups is one of the most complicated phenomena studied by social scientists, the objective of greater understanding, which could lead to reduced conflict, remains an enticing motivation. This section is dedicated to a review of the relevant contributions in the literature that have been made to that endeavour.

Before entering into the discussion of intergroup dynamics, however, it is beneficial briefly to adopt a more general approach in order to explain why individuals actually come together in the form of groups. A review of the relevant literature makes it clear that there are two fundamentally different influences behind this phenomenon [170]. These influences can be classified as being either internal or external. In other words, people come together in groups to satisfy either some internal need specific to the social psyche of humans, as discussed in §4.1, or individuals come together to fulfil some external need necessary for survival. This external drive toward survival is of special interest. From the very beginning of human civilisation, individuals recognised that it was beneficial to function within groups, with the general purpose of gathering and protecting resources. Furthermore, most, if not all, of the resources required or desired by these groups were, and are, available in limited supply. For this reason, any form of intergroup interaction will fundamentally take place on the foundation of competition (*i.e.* unless, of course, the gathering is for the simple benefit of social pleasure) [170].

These two points of view from which the literature approaches intergroup interaction are not incidental — to truly understand such interaction one must account for both these influences. Furthermore, one must appreciate that although these influences are fundamentally different,

they certainty do not take place in isolation. On the contrary, it can be argued that groups rarely compete for resources without the social element of human psyche influencing discussion. It is even tempting to claim that the drive to gather and protect resources might be the primary motive behind group interaction, but the social influences of human psyche is the element that dictates much of the behaviour within that interaction.

4.4.1 Key developments

Although rather dated, two historic contributions to the study of intergroup interactions are noted here. The first is Le Bon's theory[4] of *The Crowd* [280]. Le Bon lived in the politically volatile French society of the early nineteenth century and, being a man who did not let opportunity slip by, took advantage of his context to propose a series of propositions about the behaviour of a crowd. Some of these propositions addressed topics such as the effect of race on human affairs, the tension between the masses and the elite of a society, and the manner in which groups shape the meaning of concepts for their members [14]. As a pioneer, Le Bon was limited by his time and, as a result, the manner in which he conducted his research is rather different from what has become the norm in modern social science. Nevertheless, Le Bon's work marks a key contribution to the study of intergroup dynamics.

The second contribution worth noting, published in the year 1907, was made by William Graham Sumner [440] who hypothesised that intergroup interaction in the context of conflict followed a predictable syndrome-like pattern, which he coined *ethnocentrism*. In essence, ethnocentrism encapsulated the mechanism by which one's own group becomes the centre of everything, to the extent that all other things are weighed and rated with one's group as a point of reference [14]. In the many decades of research that followed, Sumner's theory of ethnocentrism proved exceedingly influential and marks another key contribution to the then developing field of intergroup research [71, 287].

4.4.2 The discontinuity effect

In modern literature, the phenomenon that is probably the perfect case in point of the tendency of social dynamics to influence the interaction of competing groups is the *discontinuity effect*. The discontinuity effect states that the competitiveness between groups is disproportionally larger than the competitiveness exhibited by individuals when interacting with other individuals [485]. Numerous studies have been conducted toward understanding this effect. Insko *et al.* [227], for example, documented the discontinuity effect by comparing inter-individual conflict with intergroup conflict in the context of the well-known prisoner's dilemma game (*i.e.* a mixed-motive game in which two parties are offered a choice between playing the game following either a *competitive* or a *cooperative* strategy). When the game was played by three autonomous individuals, a total of 7.5% of the responses were competitive. When the game was played by two interacting triads, however, a significant 36.2% of results were competitive. These experiments have been repeated in several studies and the results are remarkably consistent [485].

A potential contributor to the reality of the discontinuity effect that has been established in the literature is called the *diffusion of responsibility* [170], a phenomenon which can be defined as the tendency to feel less accountable for decisions made within the context of a group [170]. Meier and Hinsz [313] were the first to make the connection between these phenomena during an experiment in which they asked participants to squirt any amount of hot sauce onto a plate

[4] Many social scientists mark this publication as the beginning of intergroup studies.

that other participants would then have to eat. In some instances participants functioned within the context of a group, assigning hot sauce to either another group or an individual participant. In other instances, however, participants acted individually, again assigning hot sauce to either a group or an individual participant. In short, the study confirmed the discontinuity effect with groups consistently allocating larger quantities of hot sauce to other participants. Meier *et al.* [313] concluded the study by suggesting that a diffusion of responsibility was partially responsible for participants assigning larger quantities of hot sauce when functioning in the context of a group.

4.4.3 Power and dominion

One cannot discuss the intricacies of intergroup relations without briefly touching upon the influence of the struggle for power. In the introduction of §4.4 it was stated that one of the main drivers toward functioning within groups was originally the desire to gather and protect resources, and rightly so. This existence of groups, however, introduced a new currency — power. As Rouhana *et al.* [390] put it, groups do not only attempt to monopolise scarce resources, they also strive to gain control over other groups' resources, land and people. In fact, the very definition of power suggests a separation, a divide between individuals on the basis of association. This phenomenon is mentioned here only in passing and will be addressed further in subsequent chapters.

4.4.4 The ingroup-outgroup bias

The sociologist, William Graham Sumner believed that humans are, by their very nature, a species that joins together in groups [440]. With equal conviction, he also noted a second, equally powerful human tendency — favouring one's own group above all others [440]. In Sumner's own words, "Each group nourishes its own pride and vanity, boasts itself superior, exalts its own divinities, and looks with contempt on outsiders" [440]. This tendency was called the *ingroup-outgroup bias* [170]. As a side-note, Sumner believed that the ingroup-outgroup bias went hand-in-hand with ethnocentrism, a phenomenon discussed in §4.4.1.

Interestingly, the ingroup-outgroup bias is really two biases combined into one: The selective favouring of one's ingroup, its members and products, and the derogation of the outgroup, its members and products. These biases, however, rarely function independently and are, therefore, defined as one. The extent to which these biases influence behaviour depends on a host of situational factors, including the level of ambiguity with respect to each group's characteristics, and the level of group members' identification with the group [170]. Overall, the subtlety of the ingroup-outgroup bias nevertheless renders it robust [160, 214].

After combing through the literature, it became clear to the author that several context-specific cases of the ingroup-outgroup bias are documented as if they are psychological dynamics in their own right. It may seem unnecessary to document such cases, but a brief review is beneficial as it effectively illustrates the practical impact of the ingroup-outgroup bias. The first such dynamic is the existence of *double-standard thinking* [170], a mechanism that employs the human tendency to judge the intentions of others. Group members typically judge their own group's actions as fair and just, while condemning similar conduct by another group as unfair and unjust [135]. One group calls their warnings *requests*, while the other group calls them *threats*. A group may consider themselves *courageous* for pushing an agenda, but may be viewed as being *stubborn* by other groups [135].

A classic example of this can be found in the way Palestine and Israel view one another with respect to the major Middle East wars of 1948, 1956, 1967 and 1973. In an article titled *Misperception in the Arab-Israeli Conflict*, the psychologist RK White described how he discovered that both Palestine and Israel believe that the other side had been the aggressor in all four wars [481]. In two of these wars (the wars of 1956 and 1967), Palestine believed that Israel had attacked without any incitement. In the remaining two wars, Palestine admitted to having initiated hostilities, but claimed they did so only because they were forced to by Israel's expansionist politics [170]. Conversely, Israel felt that the first and final war of 1948 and 1973, respectively, were both examples of blatant, unqualified Palestinian aggression, while their actions in 1956 and 1967 had indirectly been caused by threats received from Palestine. Clearly, there are considerable polarities in the perceptions of these opponents [481]. This tendency to utilise a separate sliding scale when judging the actions of one's own and other groups constitutes the first more specific case of the ingroup-outgroup bias.

The next documented case of the ingroup-outgroup bias can be described by stating that when individuals categorise (*i.e.* judge) others, their perceptions are influenced more by their own predefined, category-based expectations than by what the available evidence suggests about an individual [417]. Practically, this eludes to the fact that individuals view the members of their ingroup with a certain measure of individuality, but perceive the members of an outgroup as being extensively homogeneous. This inclination is coined the *group homogeneity bias* [291] and, similar to the biases that were discussed in Chapter 3, sets the stage for a number of errors in judgement. The first of these is called the *group attribution error*, a phenomenon by which individuals generalise the characteristics of a small number of people to a larger group of which those individuals are perceived to be representative [369]. A classic example of this is the way travellers are able to deem an entire country discourteous after being treated rudely by a single native [369]. Named by Pettigrew [359], another such error is called the *ultimate attribution error*, which prompts individuals to attribute the actions of other individuals to their characters, rather than to the constraints of a situation [213].

It is important to take note of the natural tendencies documented in this section, for their influence is certainly greatly aggravated by the freedom of the subconscious. In some ways the ingroup-outgroup bias is naturally human, but its influence should remain appropriate. History tells many devastating tales of cruelty enacted by one society upon another and one wonders whether these tales would have ended differently had the actors been aware of the mechanisms dictating their behaviour.

4.5 Chapter summary

It would seem that when individuals come together to form groups, the result, more often than not, lies on the extreme ends of the spectrum between glory and disaster. The main question is: What pre-empts these extremes? This chapter was dedicated to a review of the attempts by academics to answer this question. The academic literature eludes to the fact that researchers have approached this question on one of three different levels of abstraction, namely the dynamics of the individual, the dynamics of the group, and the dynamics between groups. These levels constituted the foundation upon which the structure of this chapter was based, with §4.1 and §4.2 dedicated to a review of the dynamics of the individual, while §4.3 and §4.4 were dedicated to a review of the dynamics within and between groups, respectively.

A central theme in this chapter was that group interaction and conflict go hand in hand — conflict within the needs of an individual, but more notably, the conflict that exists among

interacting groups. Mark Twain eluded to this conflict when he claimed that "the very ink with which history is written is merely fluid prejudice" [422]. It was illustrated in §4.4 that one of the driving forces behind this conflict is the naturally human tendency to protect what is experienced as an extension of the self, namely, the ingroup. In conclusion, it is acknowledged that although this phenomenon is naturally human, its employment as a starting point for the derogation of others by opinion or deed should by no means be excused because of its natural origin. In order to progress toward the considerable potential of group performance, it is necessary to educate our tendencies. Such change starts with a certain measure of awareness. And awareness is invited by asking the right questions.

CHAPTER 5

Debiasing methods

Contents

It was Baruch Fischhoff, father of the discipline known as *debiasing*, who suggested that whenever a behavioural phenomenon is identified in some experimental context, it is nothing other than appropriate to question the robustness of that phenomenon, thereby determining the limits of its governance [159]. The preceding chapters were dedicated to a review of the academic literature related to cognitive and social phenomena that tend to govern the process of human judgment and decision making. In the spirit of Fischoff's proposition, this chapter is dedicated to a review of the academic literature related to the discipline of debiasing such phenomena. Accordingly, §5.1 is dedicated to an introduction to the field, §5.2–§5.4 are devoted to a review of dominant debiasing taxonomies contained therein, and §5.5 contains a review of the literature related to studies aimed at investigating the effectiveness of debiasing methods in practice. The chapter finally comes to a close with a short review in §5.6.

5.1 The methods of debiasing

If the study of cognitive and social bias represents an attempt at clarifying the sources and limits of the apparent wisdom of human judgement, then the study of debiasing represents an attempt at clarifying the sources and limits of its apparent folly [159]. The discipline of debiasing originated formally around 1981 when Fischhoff published an article, titled *Debiasing* [159], in an attempt to delineate how social scientists could go about conducting bias-proof experiments. He did so by presenting a typology of methods dedicated to that task. Since this publication by Fischhoff, theorists have continued to propose alternative typologies for distinguishing between the various methods used in the field of debiasing. Although a number of these methods are irrelevant to the work presented in this dissertation, either because they cannot be generalised to the context of this work or because they simply fail to hold water [429], a study of the literature uncovered three typologies that seem to dominate existing taxonomies of debiasing methods.

These are the original typology proposed by Fischhoff, a typology proposed by Larrick [278], and a typology proposed by Soll *et al.* [429]. In the remainder of this chapter, the aforementioned typologies are reveiwed, in the order listed.

5.2 A typology proposed by Fischhoff

Although the implications of Fischhoff's work have been generalised to a wider range of applications, the discussion thereof in this section is conducted in respect of the context for which it was originally intended — the task of experimental design. In this context, Fischhoff elected to distinguish between debiasing methods according to whether their working presumes that decision biases are a symptom of (1) the task executed during the experiment (*i.e.* a faulty task), (2) the nature of a judge participating in the experiment (*i.e.* a faulty judge), or some mismatch between the two [159].

5.2.1 A faulty task

In order to describe the features related to the design of a faulty task, Fischhoff distinguished among two such task-based errors: *Unfair experimentation*, in which case the task contained in an experiment is unfairly designed, and *dubious experimentation*, in which case the task contained in an experiment is easily misunderstood [159]. Based upon this distinction, Fischhoff proceeded to delineate the bias-inducing features related thereto, and the manner in which such features may be mitigated.

Unfair experimentation

In respect of debiasing unfairly designed experiments, Fischhoff [159] offered as a first line of investigation the identification of what he called *putative methodological artefacts* (*i.e.* bias-inducing features that result from the unfair design of an experiment). In order to guide an analyst's search for such artefacts, the following list of suspect features was provided:

(a) Subjects may have failed to care enough about the execution of the experimental task,

(b) subjects may have been confused by the task,

(c) subjects may not have been convinced by the experimenter's assertions about the nature of the experimental task and perceived a *pay-off structure* other than that intended by the experimenter,

(d) subjects may have been unable to express what they know, or

(e) subjects may have been asked too many questions and thus developed stereotypical response patterns to help them get through the task.

Fischhoff did not stop at the identification of putative methodological artefacts, but proceeded to delineate the following artefact-specific strategies according to which they could be mitigated: In the case of (a), analysts should raise the stakes accruing to good performance. In the case of (b), it should be sufficient to provide more careful instruction and/or more familiar stimuli. In the case of (c), analysts are to provide subjects with the assurance that their best guess at the right answer is all that is of interest and that they should respond accordingly. In the case of (d), analysts should attempt to use more familiar and/or more pliable response modes. Finally, in the case of (e), analysts should simply ask fewer questions [159].

Dubious experimentation

Interestingly, Fischhoff stated that while artefact studies (described above) carry an "implicit aspersion of experimental malpractice" (*i.e.* the allegation that the experimenter should have known better), an experimenter's failure to understand respondents' phenomenological or conceptual universe is far less condemnable [159]. Essentially, the occurrence of dubious experimentation is a product of features related to the execution of an experiment that are often difficult to control and thus their mitigation is a rather tricky task. In an attempt to address this kind of error, Fischhoff [159] presented four dynamics related to their occurrence:

(a) Participants may be pursuing additional unsolicited goals that require sub-optimal performance in the examined task,

(b) participants may possess alternative definitions for key terms other than that held by the experimenter,

(c) the execution of the experimental task may depend upon participants fortuitously making assumptions that accidentally concur with those made by the experimenter, or

(d) an experimenter may simply be insensitive to courses of action selected by participants that are, in fact, reasonable.

Interestingly, Fischhoff chose not to provide case-specific strategies for the mitigation of the aforementioned phenomena, instead simply proposing that analysts identify and then demonstrate their existence[1].

5.2.2 A faulty judge

According to Fischhoff, if the quality of an experimental task has been verified, and the bias remains, the suspicion of culpability then falls upon the respondent. This would imply one of two scenarios: (1) Participants are performing suboptimally because of the sheer difficulty of perfecting human performance, or (2) the limitations of human judgement are rendering participants fundamentally incapable of executing the experimental task. Fischhoff called these errors *perfectible judges* and *incorrigible judges*, respectively [159]. The strategies he suggested toward their mitigation are discussed below.

Perfectible judges

In order to address the issue of perfectible judges, Fischhoff adopted what is described as an *escalation approach*. The approach is considered escalating in the sense that it involves proposing a set of bias-reducing strategies, where each successive strategy reflects an increasing pessimism about the possibility of perfecting human performance in the context of experimentation [159]. In other words, if the first strategy fails, proceed to employ the (more radical) one proposed thereafter. The following strategies were incorporated into the approach:

(a) Warn participants about the possibility of bias without specifying its nature,

[1] Fischhoff seemingly assumed that the identification and subsequent demonstration of the listed phenomena will naturally result in their mitigation.

(b) describe the direction (and perhaps the extent) of the bias that is typically observed[2],

(c) provide a dose of feedback, thereby personalising the implications of the prior warning, or

(d) offer participants extended training with feedback, coaching, and any other form of teaching that may afford the respondents cognitive mastery of the experimental task[3].

This kind of approach to debiasing differs fundamentally from that suggested in respect of faulty tasks in that it is based upon the assumption that the error that has occurred is systematic and that respondents need specific instruction, as opposed to merely a fair chance, in order to succeed.

Incorrigible judges

At some point during experimentation, an analyst may conclude that a participant's successful execution of the experimental task is impossible, or rather, that it is only attainable *via* procedures that essentially coerce the participant to respond properly. In such a situation, Fischhoff [159] considered one of three responses appropriate:

(a) Replace the people participating in the experiment with some superior answering device,

(b) recalibrate fallible judgements (assuming that the amount and direction of errors are predictable), or

(c) acknowledge the imprecision in people's judgement and plan actions based on them.

Although this kind of conclusion to scientific experimentation is certainly undesirable, Fischhoff wrote that, "the decision maker or decision analyst who has given up on people in any of the above ways may still contribute to our understanding of judgement by assessing the size, prevalence, and resilience of such indelible biases" [159].

5.2.3 A mismatch between judge and task

The final kind of experimental bias identified by Fischhoff results neither from the design of the task contained therein, nor from the nature of the judge participating therein, but from the mismatch of these two elements. In other words, despite possessing all requisite skills, the manner in which a participant interacts with a task causes that participant to fail in actually putting his or her skills to use [159]. According to Fischhoff, successfully eliminating this kind of error entails making the experimental person-task system as compatible as possible which, in turn, requires either *restructuring* the experimental task, or *replacing* the current group of participants [159].

[2] In contrast to Strategy (a) and Strategy (b) proposed to supplement perfectible judges, Babcock and Loewenstein [38] illustrated that simply knowing about a potential bias is not enough to result in its mitigation. In order to mitigate the effect of a certain bias, participants had to not only be aware of its potential influence, but also understand the underlying mechanisms according to which it operates.

[3] Although the benefit of educating participants in normative rules, such as statistical principles, seems rather intuitive, several theorists have painted a slightly less optimistic picture. Nisbett *et al.* [346] and Fong *et al.* [166] both found that the benefit of teaching individuals how to use statistical rules, for example, is limited to the domain in which such rules were learnt and that they failed to have a general effect on the quality of individual decision making.

Restructuring the experiment

With respect to the former of these approaches, Fischhoff argued that as a mechanically sound aircraft requires sufficient instrument design in order to be flyable, an honest experimental task will only become tractable once it has been structured so that it actually allows respondents to put their cognitive skills to good use [159]. After acknowledging that addressing this form of bias is often context specific, Fischhoff [159] suggested the following general approaches to its mitigation:

(a) Force respondents to express what they know explicitly (instead of letting their thinking remain "in the head"),

(b) encourage respondents to search for discrepant evidence, rather than only focusing on details that corroborate a preferred answer,

(c) offer ways to decompose the experimental task into more tractable components,

(d) get the respondent to consider alternative situations related to the one under consideration in order to better understand the situation at hand, or

(e) present the participant with alternative formulations of the presented problem.

Replacing participants

In contrast to the strategies suggested above, the latter approach toward addressing a mismatch between judge and task pays no attention to the experimental task, and simply entails replacing the current group of individuals participating therein. The thinking behind this strategy is essentially that although the current group of participants are failing to complete the experimental task, it would, in fact, be possible for another group of participants to do so. In response to this scenario, Fischhoff [159] considered the following responses appropriate:

(a) Use experts who, along with their substantive knowledge, have acquired some special capabilities in processing information under conditions of uncertainty, or

(b) use a new breed of individual, one that has been educated from an early age to think probabilistically.

As this strategy may seem similar to that suggested in response to the case of incorrigible judges (*i.e.* train the participants), it is worth noting that Fischhoff distinguished between these responses in that the education of participants from a young age (suggested above) focuses on developing general capabilities, while the training of individuals focuses upon the development of a specific skill set [159].

5.3 A typology proposed by Larrick

Interestingly, the typology proposed by Larrick [278] was published under the same title as the one proposed by Fischhoff [159] (*i.e. Debiasing*). Other than that, however, the typologies are quite distinct. Unlike Fischhoff, Larrick neglected to clarify the underlying feature that governs the distinction of dissimilar debiasing approaches in the typology he proposed. He simply went about introducing the following three strategy classes: *Motivational strategies*, *cognitive*

strategies and *technological strategies*. To be precise, it should be noted that Larrick did, in fact, distinguish between these strategy classes by stating that while motivational debiasing strategies are based on the assumption that individuals possess normative decision making strategies (and will use them if the benefit of doing so outweighs the cost), cognitive and technological strategies are based on the assumption that an individual's intuitive judgment is imperfect, and need to be supplemented by alternative strategies that approach normative standards. These qualities are, however, not used to distinguish formally between the aforementioned strategy classes, they are simply mentioned in a passing fashion [278]. In order to clarify how these strategy classes differ, they are further discussed in the sections that follow.

5.3.1 Motivational strategies

Although not formally stated, motivational debiasing strategies seem to represent any attempt at debiasing that entails altering the level of an individual's "drive" to perform well. Larrick defined two subclasses into which such strategies may be classified, namely *incentives* and *accountability* [278].

Incentives

As the name suggests, incentive-based debiasing strategies involve raising the monetary stakes associated with good performance — the assumption being that higher stakes will cause individuals to spend more energy on reflection and calculation [278]. Interestingly, after introducing this class of strategy, Larrick proceeded to argue that such strategies are generally ineffective in the task of debiasing. To support his argument, he referred to the assumption that seems to underlie motivational debiasing strategies (*i.e.* that individuals do, in fact, possess effective decision making strategies, but simply fail to put them to good use because they lack the motivation to do so [278]). Larrick argued that, more often than not, this assumption fails to hold water[4]. In fact, he quoted Camerer and Hogarth [85] who stated, in an oft-cited article titled *The effects of financial incentives in experiments: A review and capital-labour-production framework*, that "there is no replicated study in which a theory of rational choice was rejected at low stakes in favour of a well-specified behavioural alternative, and accepted at high stakes."

Accountability

The second class of motivational debiasing strategies (*i.e.* accountability) is in many ways similar to the first. The only difference, really, is that the motivation-soliciting promise of potential benefit takes the form of social status instead of monetary incentive [278]. According to Larrick, the principal mechanism according to which accountability may improve an individual's decision making is that of *pre-emptive self-criticism*. The idea being that individuals will, in preparation of justifying their decisions to others, anticipate and identify the flaws in their own arguments, and thereby improve their decision making process and outcomes [278]. Like incentive-based debiasing strategies, however, this effect has been found to be limited primarily to tasks for which individuals do, in fact, possess appropriate decision making strategies and are simply

[4]Despite this general warning against the use of incentives as debiasing strategy, Larrick proceeded later in the discussion to delineate a limited number of studies that did, in fact, find certain kinds of decision making tasks to benefit from incentive-based debiasing strategies. The most recent study listed was conducted in 1995 by Stone and Ziebart [435]. These authors found that the use of incentives with respect to the execution of *multi-attribute choice* tasks lead individuals to search more extensively for information, and process that information more by *alternative* than by *attribute*, resulting in improved decision making.

failing to put them to use (see Lerner and Tetlock [285] for a review of studies related to this topic). Despite being similar to the use of incentives, the social nature of accountability as a debiasing strategy does make it unique in a number of ways. One interesting difference noted by Larrick is that accountability tends to evoke a strong need to appear consistent to others, and although rigid consistency can be detrimental to many kinds of decision making tasks, an increased measure thereof has been shown to improve tasks that involve prediction (where the inconsistent weighting of reliable cues is a major source of error) [420].

As Larrick made clear, despite being potentially useful, there are unfortunately a number of dangers associated with the indiscriminate application of accountability as a debiasing strategy. First, accountable decision makers tend to "give people what they want" (*i.e.* if an individual is aware of his or her audience's preference, the individual's decision could be biased toward that outcome [74]) [278]. Secondly, instead of motivating an individual to verify the appropriateness of his or her decisions, accountability may simply motivate him or her to construct a preselected justification for the current course of action [415].

5.3.2 Cognitive strategies

The second category of debiasing strategies in Larrick's typology is reserved for any strategy that involves altering a decision maker's cognitive approach to a decision making task. Interestingly, Larrick selected not to divide this class of debiasing strategy into additional subclasses[5], but simply proceeded to discuss three specific kinds of cognitive debiasing strategies, namely *consider the opposite*, *training in normative rules*, and *training in representation* [278].

Consider the opposite

By necessity, cognitive debiasing strategies often take the form of context-specific rules tailored to address only a limited number of biases (*e.g. the law of large numbers* [62] or *the sunk cost rule* [26]). It is thus all the more impressive that the simple strategy of "consider the opposite" has been documented to reduce several biases effectively, including the *overconfidence bias* [25], the *hindsight bias* [27] and the *anchoring effect* [337]. The strategy consists of nothing more than asking oneself the following question [278]: "What are some reasons that my initial judgement might be wrong?" As Larrick delineates, this strategy is considered to be effective because it counteracts the basic problem with oft-employed *association-based* decision making strategies — attention being paid to an overly narrow sample of evidence [278]. Despite the potential benefit of this debiasing strategy, it is worth noting that Larrick concluded his discussion thereof by warning that it should be applied with caution, as asking a individual to list "too much" counter-evidence could bias his or her decision in the opposite direction [278].

Training in normative rules

In a series of studies that took place between 1983 and 1993, a group of researchers, lead by Nisbett [345], went about exploring the debiasing potential of training individuals in the use

[5] In a paper dedicated to a review of the academic literature related to debiasing methods, Kaufmann *et al.* [266] suggested further dividing the class of cognitive strategies in Larrick's typology according to the feature such strategies attempt to alter. More specifically, Kaufmann *et al.* [266] suggested categorising this class of debiasing strategies based upon whether they entail (1) altering the *structure* of the decision-related information to suite the natural strengths of human cognition, (2) enriching an individual's self-centred perspective by getting him or her to consider the perspective of an "outside" party, or (3) drawing attention to alternative outcomes [266].

of specific decision making rules. Although not formally defined, a decision making rule seems to represent any "rule of thumb" prescribed to simplify and support a certain decision making task. The work pioneered by Nisbet produced several significant discoveries. In summary, it was illustrated (1) that the application of normative decision making rules could indeed be successfully taught (often with relatively brief training), (2) that concrete examples of abstract rules facilitated such training, and (3) that simplicity and/or familiarity with the general logic of a certain rule greatly facilitated learning (see Fong *et al.* [166], Fong and Nisbett [167], Jepson *et al.* [243], Larrick *et al.* [279] and Nisbett [283] for a review).

Training by representation

The second paradigm of training, and final cognitive strategy, discussed by Larrick (*i.e.* training by representation) seems to have been inspired by the discovery that people tend to reason more accurately when thinking in terms of frequency, compared to probability [187]. This discovery led the theorist Sedlmeier [412] to determine whether an individual's superior skill in frequency-based reasoning could be leveraged and used as a debiasing strategy. To elucidate, Sedlmeier went about conducting several studies in order to determine whether an individual could, in fact, be trained to translate probabilistic reasoning tasks into the language of frequencies, thereby improving his or her performance [412]. In summary, these studies involved testing participants on three probability-based tasks, respectively related to the *conjunction rule*, *Bayes' rule* and *conditional probabilities*. Interestingly, Sedlmeier found that participants who had been trained in frequency-based methods of reasoning outperformed those trained only in probability-based methods [412], thereby illustrating the capacity of training by representation as a debiasing technique.

5.3.3 Technological strategies

Interestingly, Larrick [278] went about introducing a number of debiasing strategies belonging to the technological class of strategies without first formally defining the manner in which this class of strategies differs from those discussed before it. But, based upon the author's interpretation of the strategies proposed to belong to this class, they seem to represent any debiasing method that involves altering some aspect of a decision making task that is external to the decision maker, in order improve the accuracy with which the task is executed. Unfortunately, the prototypical examples of such strategies that were discussed are so specific that a detailed review thereof fails to add value to the current discussion. As such, these strategies are simply listed as follows: (1) Make decisions *collectively* (*i.e.* in groups), (2) use *decision analysis* techniques such as *linear models* and *multi-attribute utility analysis* to support decision making, and (3) use *decision support systems* to supplement decision making, thereby leveraging the capabilities of computer-based decision support [278].

5.4 A typology proposed by Soll *et al.*

Although less well known than the work of Fischhoff [159] and Larrick [278], the typology proposed by Soll *et al.* [429] in an article titled *A user's guide to debiasing* has received some attention from theorists working in the field of debiasing. Like Larrick [278], Soll *et al.* [429] distinguished between debiasing methods according to the aspect of a decision making task that is the focus of debiasing. More specifically, Soll *et al.* [429] proposed that debiasing methods

be classified into one of two general categories: Those that are designed to *modify the person* partaking in a decision making task, and those that are designed to *modify the environment* in which the decision making task takes place [429].

5.4.1 Modifying the person

Soll *et al.* [429] listed three kinds of debiasing strategies to constitute the strategy class of modify the person, namely *education*, the use of *cognitive strategies*, and the use of *models to decide* [429]. Similar to the work of Fischhoff [159], Soll *et al.* listed two distinct kinds of education: Training in domain-specific decision rules, and education in general. Unfortunately, Soll *et al.* said very little more regarding these forms of education. They simply argued that although education in general has been shown to be of some value toward debiasing, the most promising form of education remains to be the adoption of domain-specific decision rules. Before discussing the next kind of debiasing method included in the general class of modifying the person (*i.e.* cognitive strategies), the following logical inconsistency between it, and the presentation of education as its own strategy class, should be noted. Per definition, educating individuals in domain-specific rules represents a kind of cognitive debiasing strategy. Yet, Soll *et al.* elected to present this form of education and the adoption of cognitive strategies (discussed below) as two separate strategy classes. In the author's view, the separation of these strategy classes constitutes an error in logic.

Moving on to the discussion of cognitive strategies, four specific debiasing strategies were proposed to constitute this class of strategy, namely *generate alternatives*, *temper optimism*, *improve judgement accuracy* and *assess uncertainty* [429]. As Soll *et al.* put it, "having a good set of alternatives from which to choose is at least as important as choosing wisely." The idea behind the generation of alternatives seems to be that the process of doing so will enable a decision maker to view a problem from alternative points of view, thereby resulting in a more informed final decision. Although Soll *et al.* [429] did not discuss the matter in detail, they suggested that the process of generating alternatives is often facilitated by doing so with respect to each objective related to the decision task individually.

The next kind of cognitive strategy (*i.e.* tempering optimism) seems to represent any attempt at counteracting the manner in which individuals tend to overestimate their chance of success with respect to their future endeavours. According to Soll *et al.* [429], this dynamic results when individuals subconsciously limit the scope of information considered with respect to some decision task to only that portion of evidence supporting some initial, preferred hypothesis (see §3.5 for a discussion on the related *optimistic bias*) [330]. In order to temper optimism, Soll *et al.* proposed a technique, initially published by Russo and Schoemaker [395], called *prospective hindsight*. The technique entails imagining having time-travelled into the future and looking back at one's prior decision making with the knowledge that the course of action previously selected had failed [429]. As several studies have illustrated, this form of contemplation often facilitates the identification of additional "causal paths" not recognised during the initial analysis, thereby resulting in a more realistic analysis of future events [324].

Naming a category within a typology dedicated to methods designed to improve decision accuracy "improving judgmental accuracy" is rather illogical, but the technique represented thereby remains of interest. The technique seems to have been inspired by the idea of the "wisdom of crowds" as studied by Galton [182] during the early 19th century. Galton found that if a large number of people were asked to participate in some judgement task (*e.g.* guessing the number of beans in a jar), the collective average of their estimates outperformed any individual's estimate [182]. Interestingly, this finding led theorists to question the effectiveness of leveraging

what they decided to call "the crowd within." The idea was that taking the average of several judgements made by an individual (separated by a time delay [471]) should outperform any single estimate [429]. The technique has indeed been illustrated to reduce bias in individual decision making [300] and, as such, forms part of the class of modifying the person, proposed by Soll *et al.* [429].

The final strategy in the class of modifying the person (*i.e.* assess uncertainty) is rather uninvolved. It essentially entails refining the manner in which a decision task is structured so that its structure facilitates the execution thereof[6]. The technique is portrayed in terms of two kinds of judgement tasks: *Interval judgements* (*e.g.* estimating with 80% confidence that a certain house will sell for between 250 and 275 thousand dollars) and *time series forecasting* (*e.g.* estimating that the value of gold will be 40 dollars per gram in 3 months' time). As Soll *et al.* [429] delineated, assessing uncertainty in terms of interval judgements would entail asking an individual to provide not one, but several, interval estimates, each related to a different level of confidence. With respect to the above example one could, for example, ask an individual first to estimate the interval that will contain the price of the house in question with a 10% level of confidence, then for a 50% level of confidence, and finally for a 80% level of confidence (which is the only level that is truly of interest). On the other hand, assessing uncertainty with respect to time series forecasting would entail progressively increasing the period over which individuals are expected to make estimates. In terms of the above example, one could, for example, ask an individual to first estimate the price of gold in one months' time, then in two months' time, and finally in three months' time (the only period that is actually of interest) [429]. The reasoning behind this kind of strategy seems to be similar to that of the generation of alternatives (discussed above) — helping individuals to consider a broader range of evidence when making decisions.

After reading the discussion of the final kind of cognitive strategy presented by Soll *et al.* [429] (*i.e.* the use of models to decide), it was unclear to the author why Soll *et al.* elected to classify this class of strategy as belonging to the more general class of modifying the person. To elucidate, the use of models to decide, as presented by Soll *et al.*, is very similar to the use of decision analysis in the typology proposed by Larrick [278] (see §5.3.3) — it entails using mathematical modelling techniques to inform decision making. Nevertheless, Soll *et al.* did indeed (for some unknown reason perhaps) characterise the use of models to decide as a modification to the decision maker and, as such, it is mentioned here.

5.4.2 Modifying the environment

Soll *et al.* [429] portrayed the modification of the environment within two main strategy classes, namely the provision of *incentives* and selection of an appropriate *choice architecture*. A discussion on incentives will add little value to the work reviewed in this chapter as Soll *et al.* simply rephrased Larrick's discussion thereof in the typology he proposed (reviewed in §5.3.1). Accordingly, this section contains only a discussion on the matter of choice architecture.

As described by Soll *et al.* [429], the term choice architecture is related to the configuration of a decision task and refers to the manner in which decision alternatives and/or relevant information is presented to a decision maker. Soll *et al.* referred to this kind of debiasing rather elegantly by characterising it as a *nudge* in the direction of improvement. Importantly, to qualify as a nudge, the selected configuration of a decision task may neither restrict a decision maker's freedom of choice nor alter the fundamental elements that constitute the decision task (*e.g.* the purchasing price of items in a product selection task) as that would unfairly influence a decision maker's

[6] As this strategy entails altering an aspect of a decision task external to the decision maker, it is unclear why Soll *et al.* [429] presented this technique as belonging to the strategy class of modifying the person.

behaviour. Instead, a *choice architect*[7] must make use of psychological principles to influence behaviour for the good in a subtle fashion [445]. Soll *et al.* proposed four classes of nudges, namely *defaults*, *nudges that induce reflection*, *nudges that induce future-focused thinking* and *nudges that kindly shape information*.

Defaults

In countries where post-mortem organ donation is the default and citizens have to take the time to opt out of donating, donation rates are approximately 90 percentage points higher than in countries where citizens must opt in to become donors [245]. Soll *et al.* [429] cited this phenomenon as an example of the potential effectiveness of using defaults to nudge decision makers in the "right" direction (assuming that being a organ donor is the right thing to do). Defaults are argued to work because

(1) they leverage a decision maker's inertia — those who procrastinate, are preoccupied, or otherwise oblivious and are automatically included [429],

(2) they tend to establish a point of reference in a decision maker's mind that *loss aversion* makes difficult to sacrifice [244], and

(3) decision makers tend to perceive defaults as being representative of expert advice [311].

After listing these dynamics that make defaults a rather powerful debiasing strategy, Soll *et al.* [429] presented a word of caution regarding their use. The main criticism raised was that their use often results in a select group of individuals defaulting to a course of action ill-suited to their personal preference. Essentially, the one-size-fits-all quality associated with defaults has the potential to benefit many people, but at the expense of seriously harming a few [427]. Despite these drawbacks, Soll *et al.* concluded their discussion of defaults by stating that they certainly "have a place in the debiaser's toolkit."

Nudges that induce reflection

According to Soll *et al.* [429], many decision biases relate to an over-dependence upon the expertise of System 1 thinking, rather than that of System 2 thinking. As such, they posit that debiasing may be promoted by making use of nudges that induce deeper reflection (see §3.1 for a review on the working of System 1 thinking and System 2 thinking). Three such nudges are proposed, namely *planning prompts*, *planned interruptions* and *active choice*. As the name suggests, planning prompts represent nudging an individual in the direction of specifying the when, where and how of an intended action [429]. The articulation of concrete plans in this manner is argued to be effective as it represents a commitment that is both psychologically difficult to break and memorable [192]. The outcomes Soll *et al.* list to benefit from planning prompts include physical training [318], meeting deadlines [131] and getting a flu-shot when appropriate [315].

The second kind of nudge introduced to mitigate bias related to an under reliance on System 2 thinking (*i.e.* planned interruptions), entails interrupting an individual's execution of some task and thereby encourage added reflection [429]. Interrupting the consumption of pre-cooked food, for example, by storing it in several containers, has been shown to reduce mindless eating [104]. The reasoning behind planned interruptions seems to be that breaking a decision making task

[7] A *choice architect* refers to the person responsible for designing the structure of a decision task [445].

into a collection of "smaller" decisions will induce deeper reflection [429]. The final kind of nudge dedicated to inducing deeper reflection (active choice), represents an alternative to the use of defaults. To elucidate, Soll *et al.* argue that by forcing individuals to select a certain decision option, instead of simply accepting some default, improves decision making. Two examples are provided to support this argument: A study in which prescription drug users were required to make an active choice between having to pick up medication at their local pharmacy, or by home delivery, which increased home delivery rates by 35 percentage points[8] [63], and a study in which newly hired employees were required to make an active choice between whether or not to sign-up for the *401k enrolment* savings plan, increasing enrolment rates by 28 percentage points over an opt-out choice scheme [90].

Nudges that induce future focused thinking

The tendency to over-emphasise immediate gratification, despite the long-term consequences of doing so, is a well documented phenomenon related to human decision making. This tendency has been linked to problems ranging from obesity [316] to insufficient retirement savings [352]. In the remainder of this section, a series of nudges, proposed by Soll *et al.* [429], are briefly reviewed as it has been claimed that they can reduce the potential effects of such "near sightedness." The following techniques are discussed: *Choose in advance*, *precommitment* and *temptation bundling*.

Based on an in-depth investigation, there seems to be no difference between the first two techniques listed above (*i.e.* choose in advance and precommitment). Both embody the utility of prompting individuals to make, and commit to, a certain decision option in advance of having to make a decision in real time [429]. An example of applying this technique would be selecting to put one's savings in a *commitment account*[9], thereby encumbering one's ability to overspend in future [35]. According to Soll *et al.* [429], this kind of technique is particularly valuable in the context of self-control problems that pit an individual's long-term interests against his or her short-term desires. The final kind of nudge proposed to belong to this category of debiasing methods (*temptation bundling*) is again similar to those discussed before it. It is, in a way, a compromised kind of precommitment. To elucidate, temptation bundling entails a precommitment to engaging in instantly gratifying activities only if they are coupled with behaviours that have long-term benefits [429]. An example of temptation bundling would be the commitment to pair eating sweet foods with physical exercise.

Nudges that kindly shape information

The reasoning behind the final kind of debiasing technique is that people tend to make better decisions when they have access to the "right" information which is, furthermore, packaged in an intuitively comprehensible and compelling format [429]. Several techniques have been proposed that build upon this proposition, namely *transform the scale*, *frame messages appropriately* and *use kind representations for probabilities*.

To illustrate the logic underlying the transformation of the scale according to which information is presented, imagine having to select between the purchase of two equally priced vehicles, namely vehicle A and vehicle B. The following information is provided: Vehicle A is considered a more fashionable vehicle, but in terms of fuel efficiency, runs 1 mile per gallon less than Vehicle B.

[8] Individuals participating in the study were offered a discount if they selected home delivery, which is rather odd as such an incentive may cause bias in the final results.

[9] A commitment account is a special kind of account that only allows the account holder to withdraw money under certain predetermined conditions.

In this scenario, one may be tempted to neglect the "slight" loss in terms of fuel efficiency (*i.e.* 1 mile per gallon), for the sake of owning a more fashionable vehicle. Conversely, consider an individual's response if the fuel efficiency of the respective vehicles were presented according to an alternative scale: Gallons per hundred miles. Using this scale, Vehicle A's fuel efficiency is a total of 17 gallons per hundred miles less efficient than Vehicle B. According to Soll *et al.* [429], altering the scale according to which numerical information is presented in this manner, and thereby exaggerating important differences between decision options, has the potential to nudge individuals onto an improved course of action.

Since the emergence of prospect theory [262], scholars have found the effect of framing upon individuals' responses to stimuli to be rather dramatic. According to Soll *et al.* [429], this calls attention to the potential of framing as a debiasing tool. Rothman and Salovey [388] illustrated this potential in the context of healthcare communication. In essence, they found that *loss-framed* messages elicited more effective illness-detection behaviour, and that *gain-framed* messages were more effective in eliciting illness-prevention behaviours (see Rothman and Updegraff [389] for more information). It follows from studies such as this, that framing may be utilised to promote certain, more appropriate responses to stimuli.

The final debiasing approach listed above concerns the misinterpretability of probabilistic information. This phenomenon is widely documented. In 2000, Hoffrage *et al.* [216] found that individuals are uniquely willing to participate in an activity when information about the occurrence of an unfortunate event (during its progression) is communicated in probability terms as compared to frequency (*e.g.* 0.01% versus 1 in 10 000). In a similar study, Galesic *et al.* [180] found that visual representation is another particularly effective way of communicating probabilistic information (*e.g.* colouring a 10×10 grid to depict the probability of some occurrence [153]). Soll *et al.* [429] drew upon these studies to suggest that this kind of representation of probability information as debiasing tool.

5.5 A collection of debiasing methods in the literature

Despite an extensive search of the academic literature, no debiasing method designed specifically for the context of interest in this dissertation could be found. Nevertheless, to illustrate that the novel debiasing method proposed subsequently does not exist in academic isolation, several such methods, tailored to alternative contexts of use, are described in the remainder of this chapter. In the order discussed, these methods relate to the task of audit judgement, legislative judgements of auditors' legal liability, fingerprint identification, and supplier selection in manufacturing.

5.5.1 Audit judgement and accountability

In 1992, Hogarth and Einhorn [217] demonstrated that belief revision (the task of re-evaluating one's judgements on a matter) consists of a sequential anchoring-and-adjustment process whereby a prior belief (*i.e.* an anchor) is revised on the basis of new information. Based upon the reality of this mechanism, they proceeded to illustrate that under certain conditions[10], a decision bias, appropriately named *recency*, starts subconsciously influencing decisions made such that information received later in a sequence is over-emphasised. More formally, recency was shown to violate the normative requirement that the net effect of encountering information A followed by information B, must equal that of B followed by A [347].

[10] The conditions in question are characterised by the complexity of a judgement task, the amount of information related thereto, and the processing strategy utilised therein [217].

Among the tasks that have proven vulnerable to the effects of recency, auditing is unique in that it has received some attention from the debiasing research community. That is, in the autumn of 1993, Kennedy [268] demonstrated the utility of a particular information processing strategy in conjunction with *accountability*, which was defined as the requirement to justify one's actions to another, in reducing the effects of recency. In essence, what Kennedy proposed is rather simple: When auditing, auditors should adopt an *end-of-sequence* strategy whereby judgements are postponed until all relevant information has been reviewed, and furthermore, they should attempt to justify their conclusions to a colleague before making them known to a client. During a discussion on why the above strategy proved to be effective, Kennedy suggested that its usefulness was in counteracting the tendency whereby auditors adopt a *step-by-step* inspection strategy according to which information and judgements are reviewed in sequence, as this was most likely what made them vulnerable to the effects of recency [268]. Notably, this explanation leaves the reader wandering to what extent the respective components of the aforementioned strategy contributed to the mitigation of recency.

5.5.2 Judges' evaluation of auditor decisions

The difficulty of adjudicating the legislative liability of auditors after the occurrence of some unforeseen, undesirable financial mishap has received considerable attention in the academic literature. The conditions that characterise the decision instance in question seemingly make it particularly vulnerable to biases associated with retrospective judgement. In 1997, Anderson *et al.* [20] contributed to the pursuit of making evaluations of legislative liability less prone to bias by publishing an article titled *The mitigation of hindsight bias in judges' evaluation of auditor decisions* in which they documented an enquiry into the utility of two debiasing techniques in mitigating the influence of a bias identified to pervade the task in question (*i.e.* the hindsight bias, see Table 3.1 for a review). The techniques considered are summarised as follows: Before making a final judgement concerning the appropriateness of an auditor's actions, a judge should consider (1) alternative outcomes that may have followed from antecedent conditions related to the audited party, and (2) alternative parties (other than the litigant) who may have been harmed if the auditor had behaved otherwise. Notably, these methods both belong to the class of debiasing strategies called *restructuring the experiment*, reviewed in §5.2.3. Somewhat surprisingly the authors found the effect of the former strategy to be statistically insignificant, while the latter method was shown to be of considerable utility [21].

5.5.3 Fingerprint identification

Forensic science was long touted as being infallable [327]. In recent years, however, academic critics have begun to realise that there may be "surprisingly little science in what is called forensic science." Notably, the science of fingerprint identification is not excluded from this accusation [377]. In an article titled *The vision in 'blind' justice: Expert perception, judgement and visual cognition in forensic pattern recognition*, Dror and Cole [137] described the case of Brandon Mayfield as exemplar of this reality. Mayfield, an attorney from Portland, Oregon, was falsely linked to a 2004 terrorist bombing in Madrid, Spain as fingerprint examiners from the Federal Bureau of Investigation determined with 100% certainty that his fingerprints matched the prints found at the scene of the bombing [137]. In reality, however, Mayfield had not even been in the country. Nevertheless, it was only after much drama (during which Mayfield spent two weeks in prison) that the Federal Bureau of Investigation was willing to admit that the prints in question, although extremely similar, did, in fact, not match [137].

In response to occurrences such as the one described above, debiasing researcher Reese [377] went about designing a collection of debiasing methods specifically aimed at facilitating the mitigation of decision biases associated with the task of fingerprint matching. She grouped these strategies into one of two categories, namely *debiasing the decision making task* or *debiasing the decision maker*[11]. A total of six debiasing strategies were proposed (evenly split between the aforementioned categories), namely *sequential unmasking*, *evidence line-ups*, *simplification*, *consider-an-alternative*, *perspective taking* and *devil's advocate*. In the remainder of this section, the respective strategies contained in each of the aforementioned categories are briefly discussed.

Sequential unmasking

Whether a result of incompetence or ignorance, as motivation to the first debiasing method listed above, Reece [377] provided evidence of the fact that fingerprint analysts are often exposed to domain-extraneous information during an examination (*i.e.* details of a crime not relevant to the task at hand). This is considered problematic as this kind of information could bias a fingerprint analyst's examination. Pinpointing this potential for error upon the decision making task, Reece proceeded to propose a debiasing technique called *sequential unmasking*. According to this technique, during both the initial and verification phases of a fingerprint analysis, all information should be *unmasked* (*i.e.* revealed) to the examiner in sequence, and only when necessary [377]. In this way, Reece [377] suggested that unnecessary exposure to potentially biasing domain-extraneous information could be prevented.

Evidence line-ups

A domain-specific instance of the confirmation bias (see §3.3 for a review), academics have shown that fingerprint examiners fall victim to what is called the *prosecutorial bias*, defined as the expectation that the prosecution's perspective on a case is, in fact, the correct one [433]. Toward mitigating the adverse effects of this phenomenon, Reece [377] proposed a debiasing method called *evidence line-ups*. In essence, the method entails introducing what are called *foils* into a fingerprint examination. That is, the fingerprints that are actually of interest are mixed with several similar-looking, but irrelevant prints [378]. The task of the fingerprint examiner is then to determine which, if any, of the prints provided match the evidence found at the scene of a crime. According to Reece [377], since an examiner would know from the structure of the evidence line-up procedure that most of the pieces of evidence are not related to the prosecutor's case, the base rate expectation that the prosecutor's take on a case will be confirmed, should be reduced[12].

Simplification

In reaching a verdict in respect of a particular fingerprint, forensic scientists are limited to one of three possible outcomes: *Identification* (representing a 100% match), *exclusion* (representing a 100% mismatch), or *inconclusive* (representing the case in which an examiner is unable to reach a decisive verdict) [377]. In other words, examiners are prohibited from making probabilistic judgements. Reece [377] considered this limitation an unfair complication of the task in question — one that exaggerates various decision biases [377]. As such, she proposed that during the initial phases of a forensic investigation, fingerprints should be partitioned into one of two

[11] Notably, these categories correspond with the debiasing typologies reviewed earlier in this chapter.

[12] Two other researchers, Risinger [378] and Miller [317], have each separately advocated the debiasing strategy in question.

categories, those that are relatively simple to identify, and those that are unduly dubious. In the remainder of the analysis, examiners should then be allowed to provide probabilistic estimates in respect of fingerprints belonging to the second category, thereby reducing the complexity, and hopefully the degree of bias, associated with the task in question.

Consider-an-alternative

The debiasing strategy called *consider-an-alternative* is equivalent to the second debiasing technique reviewed in §5.5.2, but applied in the context of fingerprint identification. Although Reece made salient that there are several use-cases for the technique in question, she provided an example of only one. That is, whenever a fingerprint analyst examining a latent print identifies what he or she believes to be a *manutiae point*[13], instead of immediately drawing a conclusion based upon that suspicion, (s)he should pause and actively considering the possibility that what has been uncovered is simply an *artefact*[14], and *vice versa*. According to Reece [377], doing so may reduce the chance of bias in decision making.

Perspective taking

The penultimate technique Reece [377] considered of utility to the task of mitigating decision bias in fingerprint matching is called *perspective taking*. As the name suggests, the technique entails attempting to consider a decision instance from the perspective of each entity related thereto before drawing any general conclusions regarding a verdict [377]. This may sound similar to the above *consider-an-alternative* technique, although perspective taking involves the consideration of an alternative point-of-view, and not an alternative scenario. Like before, Reece provided only a single example of how the technique in question may be applied. That is, if during the course of an examination, a fingerprint analyst is made aware of the fact that the prosecution expects a certain set of prints to match, thereby inducing the potential for bias, the analyst should pause, attempt to imagine that the print in question had been submitted by the defence, or even that he or she may be expected to testify on the defence's behalf, and only then proceed with the examination [377]. By doing so, the potential of falling victim to the aforementioned prosecutors bias would, according to Reece, be reduced [377].

Devil's advocate

In the final, *devil's advocate* debiasing technique, a second party, the "devil's advocate," is given the formal mandate to question and argue against a fingerprint analyst's initial conclusions [210]. Reece [377] argued that the cognitive exploration involved in this process of premeditated debate would aid analysts in uncovering inadequacies in their decision making. Notably, the devil's advocate method is similar to the aforementioned *consider-an-alternative* technique. The only difference is that in this case, the dissent process is formalised in the form of a second individual.

5.5.4 Supplier selection in manufacturing

Behavioural supply management has recently emerged as a research stream in the field of *supply chain management*. In essence, it is considered the study of how judgements made in the context

[13] A feature found in a fingerprint, such as a ridge ending, a bifurcation, or a dot [377].

[14] A random disturbance originating from dirt or from other prints that were on the surface from which the latent print was lifted [377].

of manufacturing deviate from the assumptions of the *homo economicus* school of thought [93]. In an article titled *Debiasing the supplier selection decision: A taxonomy and conceptualisation*, Kuaffmann *et al.* [266] endeavoured to document the entire collection of methods dedicated to the task of debiasing the decision of supplier selection in the context of supply chain management. They proceeded to identify a total of 68 methods, each belonging to one of five method classes, namely *decomposing*, *put yourself in the shoes of*, *draw attention to alternative outcomes*, *devil's advocate*, and *general bias awareness*. As is evident from the labels assigned to each of these categories, the methods documented by Kuaffmann *et al.* [266] are simply alternative versions of the methods that have already been reviewed in this chapter. More specifically, a discussion on these methods will be of insignificant utility to the purpose of this chapter and, consequently, such a discussion is deemed outside the scope of this chapter.

5.6 Chapter summary

This chapter was dedicated to a review of the academic literature related to the field of debiasing. It commenced in §5.1 with a brief introduction to the field in question, and this was followed in §5.2–§5.4 by a detailed review of the dominant debiasing taxonomies utilised therein. Finally, several debiasing methods similar to the one proposed later in this dissertation were briefly described in §5.5.

CHAPTER 6

The systems movement

Contents

> *"If you can't describe what you're doing as a system of things, you don't know what you're doing." — Edwards Deming*

The interconnected nature of the world is often remarkable in its subtlety, yet enduring in its presence. This chapter is dedicated to a review of the academic literature related to this interconnected nature of things and, in particular, a field of inquiry that established itself under the title of *systems theory* is of interest. This research topic is introduced in §6.1 with a short discussion on the philosophical roots of the systems movement, which is followed by a review of the general applications of the movement in §6.2. Narrowing the focus of discussion, §6.3–§6.9 contain an explication of the various paradigms that exist within the systems movement specific to the context of organisations theory, and finally, the chapter closes with a short summary in §6.10.

6.1 The origins

In many Eastern religions enlightenment follows from a recognition that everything is interconnected [234]. This sense of mutuality is, however, not unique to Eastern thinking — one also finds its origins deeply rooted within the Western philosophical tradition [234].

As with many other paradigms in the Western tradition of thinking, the first attempts at systems thinking were made by the ancient Greeks. More specifically, Aristotle first employed systems

thinking in order to elucidate the mechanisms that govern the relationship between body and soul, proposing that the soul provides the body with purpose and, consequently, its identity as an entity [393]. The eye, for example, can only see when it is connected to the rest of the body. Employing the same logic, Aristotle proceeded to transfer this notion of an organic whole to the context of politics, his conjecture being that as the body is to the eye, the state is to the individual — it shapes its purpose [393]. In general, what Aristotle claimed is that the *whole* is prior to the constituent *parts* and that the *parts* only obtain meaning within the purpose of the *whole*, and furthermore, that this principle holds true whether they are applied to the intricacies of the body and soul, or to the mechanisms of society at large [234].

Given the prominence of Aristotle within the realm of philosophy, it is unsurprising that the writings of many later philosophers (such as Plato, Spinoza, Kant, Marx and Hegel [221, 393]) are dominated by system ideas. Along with the original oeuvres of Aristotle, these writings constitute the origins of systems thinking.

As one might expect, many of the aforementioned philosophers redefined the principles of systems thinking in their own words, and as a consequence there are several somewhat unique descriptions of this paradigm. For the sake of consistency, the following general definition is adopted within the context of this dissertation: Systems thinking is founded upon the belief that the world is a conglomeration of systems which, in turn, constitute a cohesive collection of interrelated and interdependent parts, either natural or man-made. Furthermore, systems thinking holds that every system is portrayed by its spatial and temporal boundaries, surrounded and influenced by its environment, described by its structure and purpose, and expressed in its manner of functioning [232].

6.2 The systems movement

Living in the midst of interacting events, thoughts, and actions, reality as experienced from our subjective point of view, is rather complex. Humans have, however, never allowed the complexities of life to diminish their ambition toward changing the *natural* state of the world. The systems movement constitutes a notable part of this endeavour toward change — toward improvement. In fact, the scope of its application has been so extensive that bringing order to the burgeoning literature related to it, is in itself a monumental task. In order to facilitate a high-level overview of the literature on systems thinking, it is therefore beneficial to make an initial distinction between work concerned with the study of systems so as to further the theory itself and work in which systems thinking has been applied to a specific field of study (see Checkland [100]). As depicted in Figure 6.1, these categories have been labelled *general systems theory* and *applied systems theory*, respectively. As the context of this dissertation overlaps considerably with that of organisations theory, the second category is further classified into system ideas which have been applied to organisational theory and system ideas which have been applied to other disciplines (see Figure 6.1 for a list of such disciplines). The remainder of this chapter is dedicated to an exposé of systems theory as applied to organisational theory.

6.2.1 Organisational systems theory

In the *Political thought of Plato and Aristotle* [46], it is explained how Aristotle argued that in any society in which people continuously interact, divergent interests are the order of the day. For a society to remain coherent over time (*i.e.* not break up into destructive factions), those divergent interests must, therefore, be accommodated — they will not simply disappear. In the

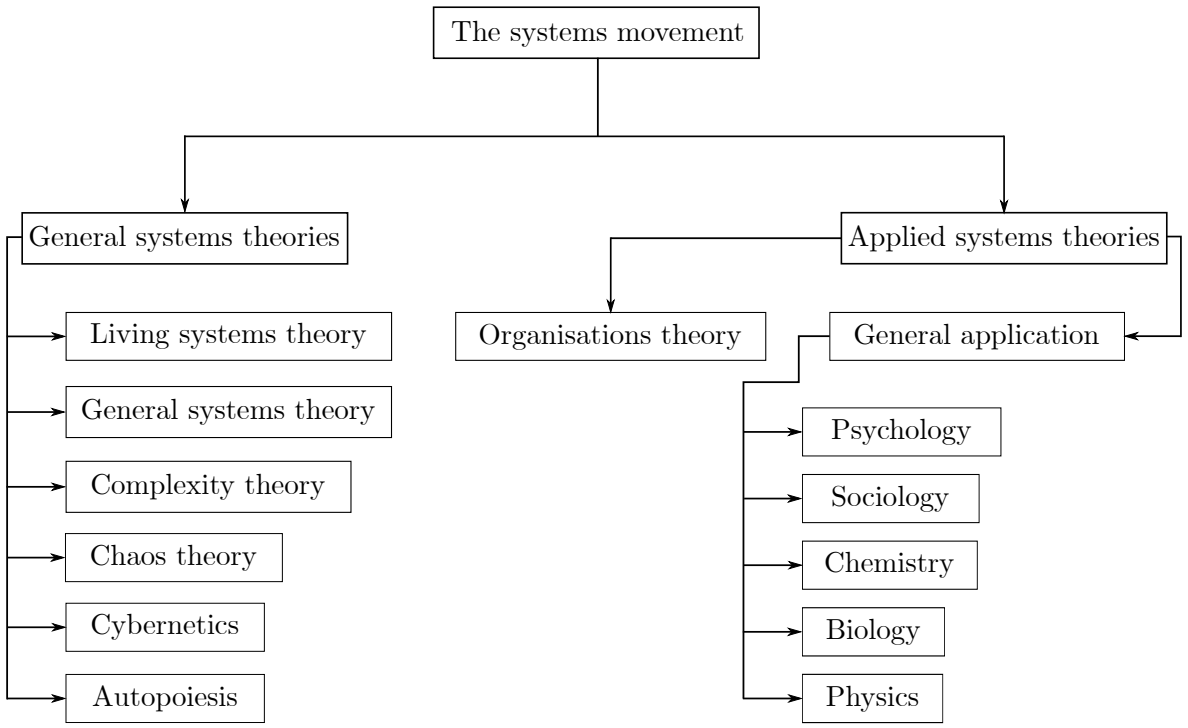


FIGURE 6.1: *A schematic representation of the systems movement, adapted from Checkland [100].*

modern world, accommodating these divergent interests is often the business of politics [254]. One of the primary mechanisms for managing such pursuits, however, originated not at the hands of politicians, but was the academic product of the economic growth that took place during the industrial revolution, a field called *organisations theory*. Organisations theory was, in a sense, the manual to the vehicle of economic pursuit of the time.

At the denouement of the revolution, three models of management were competing for precedence within this newly coined field, namely the *traditional approach*, *human relations theory* and *systems theory*. Based upon Taylor's scientific management theory, Fayol's administrative management theory and Weber's bureaucracy theory [375], the traditional approach to modelling management argued that an organisation could be likened to a machine, and modelled accordingly. This view was, however, subjected to severe critique from several commentators [232], specifically for its alleged failure to take account of the human influence within the workings of an organisation. Finding council in these critiques, theorists like Mayo, Herzberg, Maslow and McGregor endeavoured to develop a view of organisations theory that put the dynamics of humans at the centre of analysis [212]. This was the birth of human relations theory. At its introduction, human relations theory brought with it the untamed excitement that often accompanies any form of novel discovery. This excitement was, however, short-lived. By developing human relations theory in response to the shortcomings of the traditional approach, these theorists polarised their views toward the other end of the spectrum — paying attention to human needs at the expense of almost everything else [357]. Credit must, nonetheless, be given where deserved — human relations theory made salient several elements of importance in the study of organisations.

By the 1960s, the systems movement had truly made its way into the theory of organisations. A paradigm of obvious superiority, the systems approach quickly came to dominate management

theory [232]. System theorists responded to the shortcomings of preceding endeavours by criticising their limited focus upon one or two of the many aspects within an organisation necessary for optimal performance (*i.e.* the traditional approach only paid attention to task and structure, while human relations theory focused solely upon human needs). Conversely, systems theory adopted a holistic approach that considered all aspects of an organisation and recognised the fact that organisations were *open systems*, constantly interacting with their environment [234].

Within the context of organisations theory, the systems thinking paradigm produced several theories of management, each better suited for a certain context of application. Table 6.1 contains a summary of these developments, linking schools of thought to applicable methods. The remainder of this chapter is devoted to a review of the various theories of management which constitute organisational systems thinking, namely *hard systems thinking*, *cybernetics*, *machine learning*, *soft systems thinking*, *emancipatory systems thinking*, and *critical systems thinking*.

TABLE 6.1: *An explication of the theories that constitute organisational systems thinking.*

Paradigm	Methodology	Class/Method
Hard systems thinking [234]	Systems analysis [234]	Analytical modelling [486]
	Systems engineering [198]	Quality function development [326]
	Traditional operations research [115]	Analytical modelling [486]
Cybernetics [232]	Organisation cybernetics [232]	Beer's viable system model [234]
	Management cybernetics [53]	Black box technique [232]
		Negative feedback technique [232]
		Variety engineering [232]
Machine learning [88]	Supervised learning [273]	Decision trees [356]
		Support vector machines [273]
	Unsupervised learning [185]	k-Means clustering [356]
		Principle component analysis [185]
	Semi-supervised learning [96]	Kernel density estimation [356]
	Reinforcement learning [441]	Policy iteration [441]
		Value iteration [441]
Soft systems thinking [100]	General	Soft systems methodology [102]
		Social system design [113]
		Strategic assumption surfacing and testing [232]
		Social systems sciences [7]
	Soft operations research [168]	Strategic choice [178]
		Strategic options development and analysis [3]
Emancipatory systems thinking [127]	General	Critical systems heuristics [458]
		Community operations research [246]
Critical systems thinking [127]	General	Total systems intervention [234]

6.3 Hard systems thinking

It is interesting to note that the term *hard systems thinking* did not appear in the academic literature until Checkland employed it in 1975 to justify the development of soft systems thinking (to be discussed later in this chapter) [98]. This is interesting because the paradigm that the term came to represent was certainly around long before its coinage. The first strands of hard systems thinking is often traced back to a period during, and immediately after WWII — a period during which scientific research methods were being applied to the study of operational processes rather than natural phenomena.

Delving into the depths of the hard system's history of thinking is a daunting task, as the philosophical and theoretical points of view that underpin its development are often so taken-as-given that they are rarely declared explicitly. It is, however, possible to elicit such understanding by studying the similarities among the problem solving methodologies that are held to be representative of the hard systems school of thought, namely *traditional operations research*, *systems analysis*, and *systems engineering* [234]. An in-depth explication of these methodologies and the methods they contain would, however, not be a wise endeavour, as several books could be written on the methods within each. The focus in this section is, therefore, restricted to a brief review of each methodology, with the purpose of illustrating the more general hard systems school of thought underlying each. To that end, consider the following definitions:

Operations research can be defined as the application of scientific methods to complex problems arising in the pursuit and management of organised systems of people, machines, material and capital within several contexts of endeavour. The distinctive characteristic of this methodological approach is the development of scientific models of the system under consideration in order to predict and compare the outcome of several potential courses of action [233].

Sytems analysis is a manner in which to study complex problems of choice under conditions of uncertainty by eliciting the costs, risks, and utility associated with alternative courses of action. It finds application in decision aid of public and private decision makers [323].

Systems engineering is an interdisciplinary approach towards designing complex systems in such a manner that the subsystems of its composition are combined, verified, and operated in the most effective way [56].

To the trained eye, these definitions reveal various characteristics. First, it can be seen that the purpose of scientific inquiry, as employed above, differs from that with which it was traditionally associated (*i.e.* the advancement of knowledge for its own sake). Hard systems thinking finds purpose, not in the endless pursuit of new knowledge, but in the application of what is known in and about real-world instances, with the general purpose of serving the interests of decision makers [233]. A second characteristic belonging to the class of hard systems thinking relates to the proposition that models, predominantly of a mathematical nature, are able to do for the management sciences what laboratories do for the natural sciences — provide a controlled space within which the methodological study of phenomena can take place [233].

This belief in the practice of modelling exhibits several underlying assumptions inherent to hard systems thinking with respect to the state of the real world. The work of Peter Checkland [99]

comes to mind as an excellent elucidation of these assumptions. Checkland argued that there are four basic assumptions that constitute the thinking behind the hard systems approach, namely:

1. There is a present state of a system, S_0,
2. there is a desired future state of a system, S_1, which is known,
3. there are alternative ways of getting from S_0 to S_1, and
4. the analyst's role is to identify the best way of getting from S_0 to S_1.

These assumptions lie at the heart of the hard systems approach, constituting the backbone of the methodologies that represent this paradigm. By likening the purposeful, social activity of human beings to the naturally forming physical systems of nature [127], hard systems thinking employs these assumptions to argue that the world can be modelled to represent a conglomeration of subsystems, each entertaining its own clearly defined purpose through routine and repetitive operation toward realising predetermined goals [127].

Before moving on to a discussion on the next systems thinking school of thought, it is worth discussing the methods that function within the methodologies presented above. Table 6.1 contains several such methods. Although not exhaustive, these methods are considered to be relevant within the context of this section. Most of the methods that function within the hard system paradigm reside in the class of techniques called *analytical modelling*, which can be defined as the mathematical replication of the logical relationships that are believed to govern the behaviour among the components of a system under investigation [486]. Models that belong to this analytical class can be divided into four categories, depending on whether or not they model behaviour in a *discrete* or a *continuous* manner, and whether that behaviour is modelled *deterministically* (*i.e.* according to a fixed set of rules) or in a *stochastic manner* (*i.e.* according to a statistical distribution of probability). Figure 6.2 contains a schematic representation of some of these categories and a suggestion of how a limited number of modelling paradigms can be classified within them (as interpreted by the author).

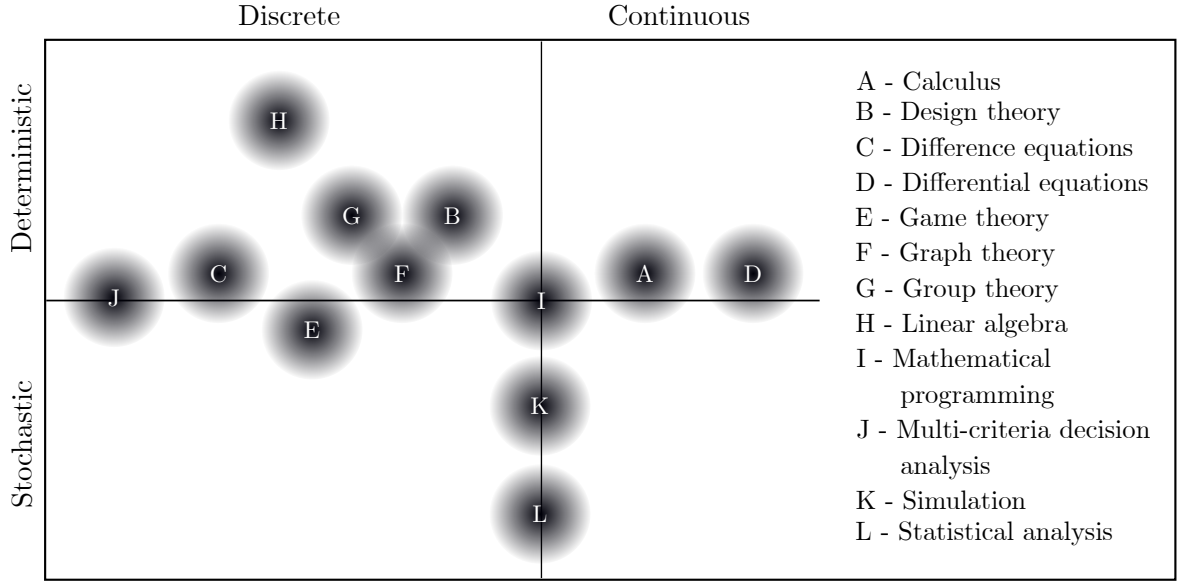


FIGURE 6.2: *Types of modelling disciplines.*

Taking another step down the staircase of classification, one can see that within each of these modelling paradigms, there are what may be called *typical problem types* (*i.e.* problems typically associated with these modelling paradigms). Table 6.2 provides an (incomplete) explication of such problem types (again as interpreted by the author).

TABLE 6.2: *Typical problem types tackled within selected modelling paradigms.*

Modelling discipline	Problem types modelled
Calculus	Inventory management problems
	Queuing theory problems
Design theory	Experimental design problems
	Sport tournament scheduling problems
Difference equations	Population dynamics models
Differential equations	Blending problems
	Control problems
Game theory	Problems of strategy formulation
Graph theory	Location problems
	Routing problems
	Timetabling problems
	Resource flow problems
	Project management problems
Group theory	Enumeration problems
Linear algebra	State transition problems
Mathematical programming	Location problems
	Routing problems
	Timetabling problems
	Resource flow problems
	Project management problems
	Lot sizing problems
	Job shop scheduling problems
	Capital budgeting problems
	Resource assignment problems
	Inventory management problems
	Transportation problems
Multi-criteria decision analysis	Preference modelling
	Trade-off analyses
Simulation	Warehousing management problems
	Production process management problems
	Resource flow problems
Statistical analysis	Regression problems
	Classification problems
	Clustering problems
	Inference modelling
	Forecasting problems

In conclusion, it can be said that early systems theory, as presented by the hard systems approach, reflects a positivist epistemology (*i.e.* the belief that observable evidence is the only form of defensible scientific discovery), with the methods therein focusing on prediction and control as inspired by the natural sciences [233]. This manner of inquiry has proven itself useful in the study of several kinds of systems, although it should be stated that the scope of its application must be selected with care. For example, when it comes to the study of complex human problems (*e.g.* poverty) the application of hard systems thinking may be inadequate [230]. To defend this statement, an assertion made by Watson *et al.* [127] is paraphrased: Through trial and error, systems researchers have concluded that social systems are too complex to be studied

by the black and white methodologies of hard systems thinking, the result more often than not being an inaccurate and unsuitable interpretation of the world. The intended purpose of this statement is not to belittle the capability of hard systems thinking as a solution methodology, but rather, to accompany its considerable potential with an appropriate level of caution.

6.4 Cybernetics

Etymologists have traced the origins of the expression *cybernetics* to the Greek word *kybernetes* — meaning the art of steermanship [232]. At the time of its coinage, *kybernetes* referred principally to the piloting of a vessel, an exceptionally dangerous occupation in those days. Around the time of Plato, philosophers expanded its applicability by employing it to refer to the act of steering the *ship of state*, a role typically assumed by the acting head of state. Throughout the history of the term *kybernetes*, the golden thread of its use has been the implied act of control, whether that be in a technical or a political sense. During his time at the Massachusetts Institute for Technology, Wiener became the first mathematician to take this notion of control and apply it to a specific field of study. His book titled *Cybernetics: Control and communication in the animal and the machine* [483], published in 1948, marked the exact moment of this transition. As the title of his book suggests, Wiener defined cybernetics as the science of control and communication [233].

Ashby, another pioneer of the cybernetic discipline, published his best-known book titled *An introduction to cybernetics* in 1956 [33]. In its initial chapters, Ashby goes about introducing a notion that became for many the "Newtonian law of management," namely the *law of requisite variety* [232]. This law states that only variety can destroy variety, or in more practical terms, if order is to remain in a system, the variety of the controlling mechanism must be at least as large as the variety of the system it is controlling [34]. Interestingly, this notion seems to coincide with a dimension of system complexity as defined by Jackson [232] (*i.e.* the number of states a system or its output can assume). This overlap, between the nature of complexity in a system and the purpose of cybernetics is, in fact, not accidental or surprising — cybernetics was developed to accommodate systems of extreme complexity, self-regulation, and probabilism [52].

Cybernetics as a discipline stretches far and wide — so wide, in fact, that its boundaries cannot properly be illuminated within the scope of this chapter. Fortunately, the basic building blocks that constitute the cybernetic methodology remain constant regardless of the context of application. As such, the remainder of this section is devoted to a review of the core principles of the cybernetic methodology as applied to the context of management (*i.e.* organisations theory). The methodology can be divided into three main building blocks, or rather, three main methods of dealing with different kinds of complexity in a system, namely the *black box technique*, the *negative feedback technique*, and *variety engineering* [233].

6.4.1 The black box technique

During its development, the proponents of management theory recognised that one can effectively manage what one can predict with a certain measure of certainty. While hard systems theorists responded to this insight by designing an approach that depended upon an in-depth understanding of a system in its entirety [234], the cybernetic school of thought recognised that it was not always necessary to understand why every part of a system behaves the way it does. In fact, cybernetic theorists argued that there are instances in which exceedingly complex components within a system *should not* be modelled in the hard systems sense of the word, and

that modelling such components would only result in an inaccurate representation of the real world [232]. In other words, complex systems are frequently more than the sum of their parts — in some instances, the parts of a system cannot be described in a manner that reproduces the pattern of behaviour exhibited by the system as a whole. As Ashby put it: "The way not to proceed in approaching an exceedingly complex system, is by analysis" [33].

The cybernetic solution to this extensive measure of complexity is to define certain parts of a system as a *black box* [232]. Then, utilising a method of input manipulation and output classification, data can be collected that are representative of the system's behaviour. These data may then be analysed to discover any regularities in behaviour that could represent the black boxes mechanisms. By doing this, an analyst is essentially developing a model of the black box, but in a very different sense than the hard systems approach would encourage. It is important to note that a black box does not necessarily have to contain a *whole* system — certain exceedingly complex subsystems within a greater system can be described by various black boxes even if the systems within which they operate are described analytically. A schematic representation of the black box technique is provided in Figure 6.3.

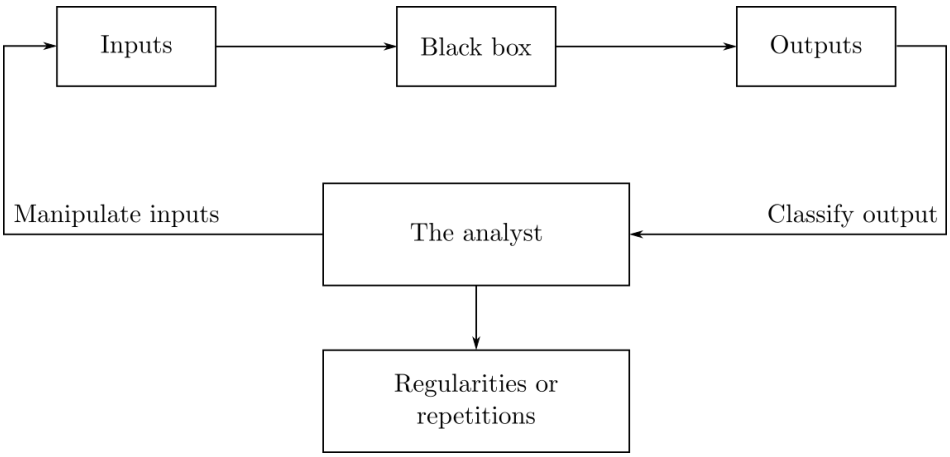


FIGURE 6.3: *The black box technique, adapted from Jackson [232].*

In conclusion, note that within the context of cybernetics the black box technique is subjective. The judgements made in respect of regularities discovered within the behaviour of a black box are based on subjective experience. It is, therefore, important that an analyst does not jump to any conclusions prematurely. A system must be studied for a sufficient period of time before attempting to describe the behaviour thereof [54].

6.4.2 Negative feedback

As discussed in the previous section, the black box technique's main purpose is to generate a set of patterns by which a system's behaviour can be predicted. The strategy of negative feedback is rather different in its purpose — instead of prediction, it focuses on control. More specifically, negative feedback is a mechanism by which a system can reach a state of self-regulation [232]. The negative feedback technique is typically employed in a situation where it is possible to model the components of a system, but the optimal strategy for managing the behaviour of the system is exceptionally difficult to identify. Essentially, negative feedback then gives a system the ability to regulate its behaviour based on observation of the real-time state of the system.

As depicted in Figure 6.4, its operation can be summarised as follows: The output of a system is continuously compared with some predetermined goal. If the system is not achieving its goal,

the margin of error (*i.e* the negative feedback) is employed as the basis for adjustments to the system, which then brings it closer to its goal [234]. This kind of closed-loop control system is exceptionally effective, since any deviation from the predetermined goal automatically sets in motion adjustments to the system aimed at bringing it back on track [404].

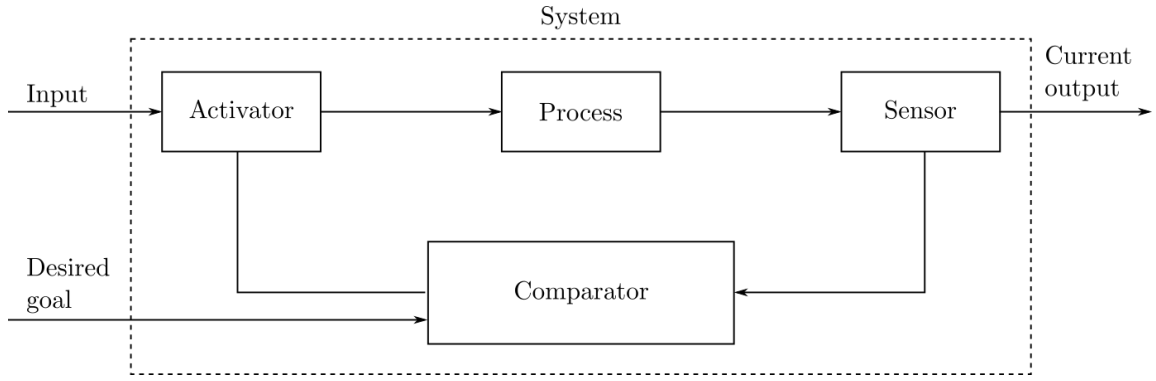


FIGURE 6.4: *The negative feedback technique, adapted from Jackson [232].*

A final point should be made before closing the discussion on negative feedback. This section has only touched on simple, first-order feedback systems [404]. More sophisticated (*i.e.* second-order and third-order) feedback systems also exist. These systems are capable of choosing among several responses to changes in a system and can even change the goal state of the system in response to feedback processes [232].

6.4.3 Variety Engineering

> *"The world is woven from billions of lives, every strand crossing every other. What we call premonition is just movement of the web. If you could attenuate to every strand of quivering data, the future would be entirely calculable. As inevitable as mathematics." — Sherlock Holmes*

The above iconic quote from the British television series *Sherlock* captures a version of reality not common to man. Managers of the 21st century are typically unable to make accurate predictions about either the organisations they manage or the environments within which they function. Consequently, organisations must develop the ability to respond to these uncertain occurrences if they are to be successful. Fortunately, Ashby held out a helping hand to managers when he coined the concept of *variety* [34], which can be defined as the number of states a system can adopt. It is, therefore, a measure of complexity [234]. As Ashby's law of requisite variety put it, only variety can destroy variety. Thus, in order to control a system's behaviour, one's response to a system needs to possess as much variety as the system itself. The cybernetic solution to this dynamic is to either reduce the variety of the system (*i.e.* variety engineering) or increase the variety of one's responses to changes in the system (*i.e.* variety amplification) [232]. This process of balancing variety constitutes the core principal of variety engineering [54].

In conclusion, the development of the field of cybernetics represented significant progress, it equipped managers with the capacity to better respond to the uncertain environment within which they function. Accordingly, cybernetics has been heralded as a unique paradigm within the systems movement.

6.5 Machine learning

The ability to learn, to develop an understanding of stimuli and respond in an appropriate manner, is an intricate, multi-faceted phenomenon. For a long time, this ability was thought to be unique to living organisms. Since the mid 20th century, however, this belief is no longer appropriate. In 1959, Arthur Samuel, a pioneer of the digital machine age coined the term *machine learning*, defining it to be the field of study that equips computers with the ability to learn without being explicitly programmed [400]. The onset of this new age, heralded a new dawn in both industrial and scientific inquiry. Interestingly, including machine learning as one of the system thinking paradigms has not been standard practice within the academic literature. By the end of this chapter, the author hopes to have illustrated why such an exclusion of the machine learning paradigm is inappropriate.

6.5.1 Machine learning meets cybernetics

There has been much debate about where machine learning, as an academic discipline, actually fits in. This confusion is believed to be due to a lack of knowledge about the somewhat lesser-known paradigm of cybernetics. To the attentive eye, however, there is an undeniable connection between machine learning and cybernetics. This connection is hinted at, but not explicitly established in the existing academic literature. It is nevertheless formally suggested here that cybernetics was the precursor of machine learning. More specifically, machine learning borrowed and built upon two concepts that were first developed within the cybernetic paradigm, namely the *black box technique* and the *negative feedback technique*. These concepts were formally defined in §6.4, but to summarise: The main reason for adopting the black-box technique was to be able to model exceedingly complex real-world systems, while the main purpose of employing the negative feedback technique was to identify a policy by which to manage exceedingly complex, dynamic real-world systems. Within the machine learning paradigm, these methods have maintained their purpose, but have adopted a new nature — the ability to learn autonomously.

Going into slightly more detail, the black box technique gave analysts the ability to model exceedingly complex systems without having to explicitly model the system under consideration (in the hard systems sense) [232]. This meant that instead of having to define the elements and relationships that governed the behaviour of a system, analysts could simply study the system's inputs and outputs as data descriptive of its nature. Within the cybernetic paradigm it was the responsibility of the analysts to study the data with a view to identifying patterns that related the inputs to the outputs. These patterns would then become a model of the real-world system, enabling analysts to predict how new inputs would be transformed into outputs [483]. This mechanism of prediction was, however, greatly limited by the cognitive capacity of the human mind to recognise appropriate patterns.

The negative feedback technique was in its own right a powerful method of coping with the uncertain environment within which most systems operate. Instead of demanding foreknowledge of the future, negative feedback gave the analyst the ability to build a measure of flexibility into a system that was more capable of dealing with uncertainty [483]. But once again, this mechanism of control was rather limited by the analyst's ability to design an appropriate measure of flexibility (*i.e.* a policy) that could respond to unexpected disturbances of a system. These limitations on the part of the analyst created a niche for the machine learning paradigm. Machine learning responded to the cognitive limitations of the human mind by stepping outside its domain, into

the domain of artificial intelligence [394]. Instead of expecting analysts to identify the patterns that model a system, or a policy that optimally manages a system, machines were given the ability to identify such patterns and uncover such responses autonomously.

6.5.2 Predominant paradigms

As the field of machine learning matured, four main paradigms of *autonomous learning* emerged. These paradigms are *supervised learning*, *unsupervised learning*, *semi-supervised learning*, and *reinforcement learning*. In the supervised learning paradigm, a machine is presented with *labelled* data corresponding to *dependent* and *independent* variables with the end goal of approximating an underlying function (*i.e.* pattern) which is best able to link input features (*i.e.* independent variables) to the output variables (*i.e.* dependent variables) [273]. Supervised and unsupervised learning are rather similar — in fact, the only difference is the feature that the data provided to the machine are *unlabelled* [65]. In other words, no guidance is given as to the underlying function that best describes the system being modelled. Semi-supervised learning is a hybrid of supervised and unsupervised learning that employs a combination of labelled and unlabelled data [96].

The paradigms of supervised, unsupervised and semi-supervised learning were (it would seem) all modelled on the black-box technique. The paradigm of reinforcement learning is, however, a method of a different kind — one that was modelled according to the negative feedback technique. Reinforcement learning is concerned with how a *software agent* (*i.e.* the machine) should take actions within an environment so as to maximise some notion of *cumulative reward* [441]. This process takes the form of a negative feedback cycle. Evidently, reinforcement learning differs from the other paradigms in its overall focus on control, compared to a specific focus on prediction within the other machine learning paradigms.

In summary, machine learning took the building blocks of the cybernetics paradigm and gave it the capacity to be applied to systems of greater complexity and scale than ever thought possible. As cybernetics built upon the foundation of hard systems thinking, machine learning built upon the foundation of cybernetics. It is not claimed here that machine learning is a subfield of cybernetics, or that the two are one and the same. Simply put, it is suggested that cybernetics was one of the foundational building blocks of the machine learning movement — a movement which has proven its capacity to facilitate the study of exceedingly complex real-world systems.

6.6 Soft systems thinking

During the mid-19th century, an active debate ensued as to the objectivity of the systems movement's problem solving techniques. Several well-known theorists, such as Rittel [382], Webber [382], Ackoff [6] and Checkland [100], later made significant contributions to this debate. Arguably, the most notable of these contributions was by Rittel and Webber in an article titled *Dilemmas in a general theory of planning*, in which they differentiated between *tame* and *wicked* problem types. The general points of critique highlighted during the debate were that the system techniques of the time were limited because of a lack of appreciation of the subjective nature of the process of problem formulation and the strict prerequisites that had to be met before being appropriately employed (*e.g.* predefined objectives) [233]. It is not claimed here that systems practitioners did not, at times, recognise the need to adapt techniques in order to accommodate real-world circumstances, but rather that the formally defined, academically documented techniques did not mirror the same sensitivity to real-world conditions.

As a result of the elucidated limitations of the systems approach of the time (*i.e.* consisting predominantly of hard systems thinking) a family of alternative methods, collectively called *soft systems thinking*, was developed with the purpose of supplementing the limitations of its predecessor [101]. The methods within this family consist of several technical procedures which are linked together by social interaction. That is, unlike the algorithmic nature of the traditional problem solving approach, soft systems thinking is largely qualitative in nature [430]. Over the years, several academics have proposed novel methods within the soft systems paradigm, some of which have stood the test of time, while some turned out to be but sifting sand under the weight of professional scrutiny.

As in previous sections, the scope of this section renders infeasible a detailed description of the methods that are dominant within the soft systems approach. Instead, in a manner pioneered by Jackson [232], the discussion is built around the aphorisms proposed by Churchman toward capturing the essence of soft systems thinking. More specifically, introducing the aforementioned aphorisms attention is afforded to a discussion on each of the three methods that are held to be representative of the soft systems approach [232], namely *soft systems methodology* (SSM), *strategic assumption surfacing and testing* (SAST), and *social systems sciences* (SSS). The section finally closes with a short review of a relatively recent field within operations research, called *soft operations research*, that was developed in response to the elucidated limitations of the hard systems paradigm within which traditional operations research functioned.

6.6.1 Churchman's social systems design

The systems movement [114], published by Churchman in 1979, was the product of several years of philosophical speculation and is truly difficult to summarise. Fortunately, in the closing chapters, Churchman described several propositions which capture the main message of the book — a message that became for many the building blocks of the soft systems philosophy. The first of these aphorisms is:

> *"The systems approach begins when first you see the world through the eyes of another."*

For Churchman, the soft systems approach begins with an understanding that the world can be viewed from several radically different perspectives that typically accompany conflicting philosophical positions. As Kant put it, no matter the view of the world one holds, it is inevitably based upon several taken-for-granted, *a priori* assumptions [265]. In this way Churchman encouraged system practitioners to embrace the fact that distinctly different evaluations of social systems can and do exist [233]. The only way to develop a wholesome understanding of a system is to study it from as many perspectives as possible [112]. Churchman supported this claim with a second aphorism:

> *"The systems approach goes on to discover that every world-view is terribly restricted."*

In *The design of inquiring systems* [113], Churchman defends the above aphorism by considering several perspectives from which five highly intelligent philosophers might describe a system — ultimately concluding that each description in isolation, is incomplete [113]. The only way to move closer to an objective understanding of a system was to combine as many subjective descriptions of a system as practically feasible. Importantly, Churchman followed this statement with the conjecture that an individual's adopted world-view is exceptionally difficult to change,

in other words, individuals find it difficult to accept the fact that their world-view is limited. In the data-rich modern state of the world, it is tempting to claim that individuals need only be exposed to information contrary to their world-view for them to acknowledge the limitations thereof. The effect of such exposure is, however, diluted as any apparent contradiction is simply interpreted, and thereby *corrected*, according to their fixed presuppositions [112]. According to Churchman, it is only through *dialectical debate* that one can truly expose decision makers to the restricted nature of their world-view.

Formally defined, the process of dialectical debate adopts a form of inquiry first employed by the German philosopher Hegel [232]: Propose a *thesis*, challenge that thesis with an *antithesis*, *synthesise* the thesis and antithesis into a new, improved thesis, and repeat the process. In the context of a dialectical debate, a thesis would be a world-view that gives meaning to a system, an antithesis would be a second relevant world-view based on entirely different assumptions, and the synthesis would be the process of considering both world-views and bringing them together in a way that produces a view of the system that is greater than the sum of its parts (*i.e.* the original world-views) [233]. This iterative process of moving closer and closer to an objective understanding of a system brings the current discussion to Churchman's final aphorism:

> *"There are no experts in the systems approach."*

This warning is one that should be taken to heart most seriously by systems practitioners themselves. Because of the human-infused state of real-world systems, the aims and objectives being pursued during problem solving will inevitably involve ethical considerations of moral judgement, an area of consideration in which there can be no expert [232].

6.6.2 Soft systems methodology

Doubtless, it is unwise to claim that any methodology of scientific inquiry has matured to the point of *perfection*. Such a claim would simply illustrate a misunderstanding of the scientific discipline. The truth, rather, is that many methods of scientific inquiry have become so taken-as-given that they seem in some sense *permanent*. To the soft systems theorist, SSM is certainly one such method [103].

Casually defined, the SSM can be divided into four phases, with each phase connected to every other phase, as depicted in Figure 6.5. This interconnectivity is not accidental — its purpose is to illustrate that the activities of the SSM process do not constitute a once-off sequence of steps. For example, employing the *activity models* of phase 2 during phase 3 will ultimately lead to the discovery of new problem elements — deeming a return to phases 1 and 2 necessary [102]. Thus, SSM is fundamentally an iterative process. The father of SSM, Peter Checkland, developed it in this manner so that it would be able to accommodate the unique, continually changing, conflict-laden nature of human interaction [103]. For a complete discussion on the SSM, the reader is referred to Checkland [102].

6.6.3 Strategic assumption surfacing and testing

A methodology for approaching ill-structured problems, SAST can be applied as a dialectical approach to policy and planning problems. It was designed to facilitate an exploration of the underlying assumptions that constitute the *status-quo* within a given context. SAST is typically employed when the formulation of a problem is of a messy nature [232]. Its four core principles, namely *adversarial debate*, *participation*, *integration* and *managerial-minded support* can be

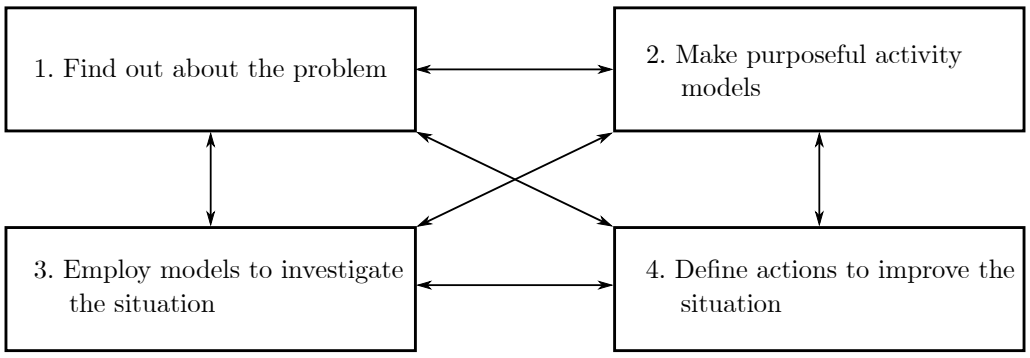


FIGURE 6.5: *A schematic representation of the SSM process.*

summarised by claiming that ill-structured problems are best understood when the conflicting perspectives of as many individuals as are affected by a problem are integrated into a higher-order synthesis which, in turn, deepens the managerial understanding of the organisation [307]. The four basic steps that constitute the SAST methodology are summarised in Figure 6.6. For a complete explication of the SAST methodology, see Mason and Mitroff [307].

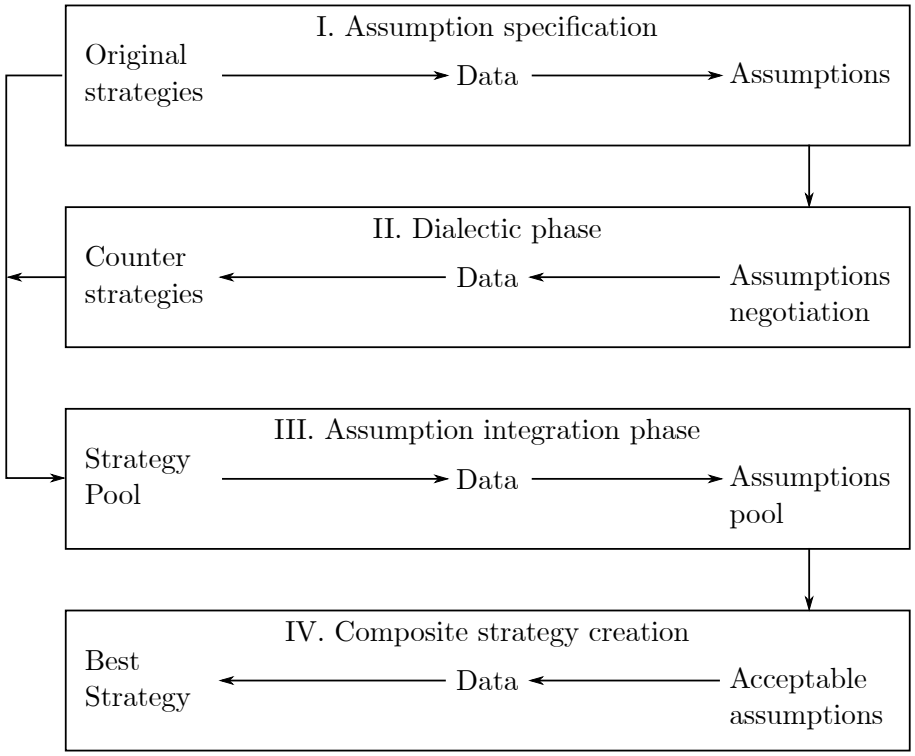


FIGURE 6.6: *A schematic representation of the SAST process, adapted from Mason and Mitroff [307].*

6.6.4 Social systems sciences

In like manner to Churchman, Ackoff [8] rejected the notion that objectivity can be obtained by constructing value-free models of a system. He believed that the only way a measure of objectivity in opinion can be obtained is by the open interaction of individual subjectivities [234]. This belief gave rise to the formalisation of a process that became the central operating tool of Ackhoff's SSS, called *interactive planning* [232].

Interactive planning is built upon three principles that are incorporated into five phases of activity [8]. The first principle, namely *the participative principle*, speaks to the importance of including relevant stakeholders throughout the planning process. The second principle is that of *continuity*, which emphasises the fact that organisations should be able to adapt in harmony with the ever-changing environment within which they function. The final principle, *the holistic principle*, warns analysts to not lose sight of the big picture, and to include all levels of activity within their planning [232]. Ackoff took these three principles and made them practical in the form of five phases of activity, schematically represented in Figure 6.7. These phases constitute a continuous, iterative process — they are never completed, only abandoned. For a complete explication of the SSS methodology, the reader is referred to Ackoff [8].

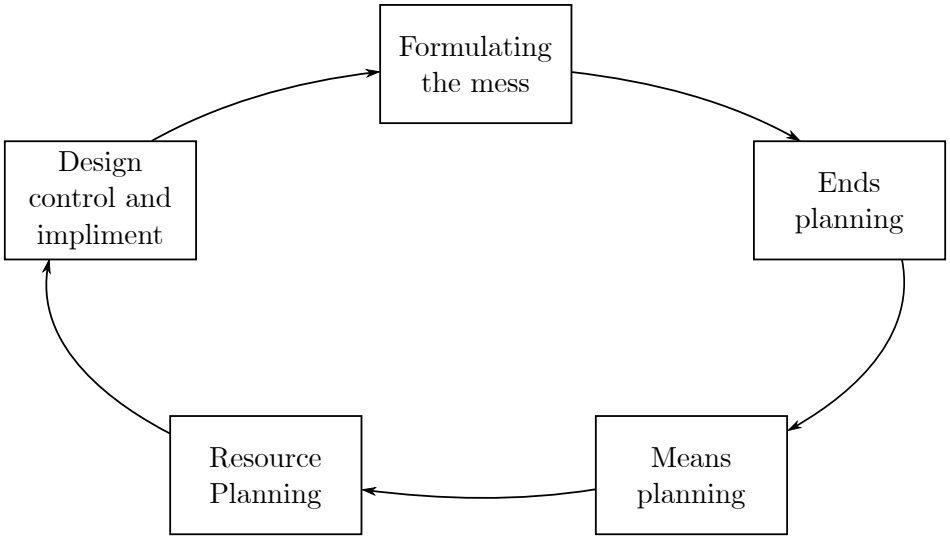


FIGURE 6.7: *A schematic representation of the SSS process, adapted from Jackson [232].*

6.6.5 Soft operations research

One of the most interesting developments to take place as a result of the elucidated shortcomings of the hard systems approach was the development of a field which came to be known as soft operations research [232]. In many ways, the development of soft operations research was an attempt at bridging the gap between the methods of the traditional systems paradigm and the intricacies of real-world problem solving by combining the analytical advantages of the traditional approach with the soft paradigm's priority of accurately capturing and structuring the intricacies of a problematic situation. There are two approaches to structuring problems in this manner which tend to dominate the realm of soft operations research, namely the *strategic choice approach* (SCA), and *strategic options development and analysis* (SODA). In the remainder of this section, these approaches, as well as a third approach which is of interest, are briefly discussed.

Strategic choice approach

When Friend and Hickling wrote *Planning under pressure: The strategic approach* [178], it was the first methodically mature exposition of the strategic choice approach. In the years since, the approach has gathered support among many operations research practitioners [177]. At its core, the strategic choice approach is dedicated to the management of uncertainty in a strategic

manner [319]. As depicted in Figure 6.8, the SCA can be divided into four main modes of thinking, namely *shaping*, *designing*, *comparing* and *choosing*. These modes of thinking are concerned with the shaping of problems, the designing of possible courses of action, a set of techniques for comparing those courses of action, and finally a set of techniques for choosing a course of action, respectively. On a lower level of operation, each of these modes of thinking can be broken down further into several phases of purposeful activity. These activities are designed to respond to three different kinds of recognised uncertainties, namely uncertainties within the *working environment*, uncertainties pertaining to *guiding values*, and uncertainties related to the *decision space* [234].

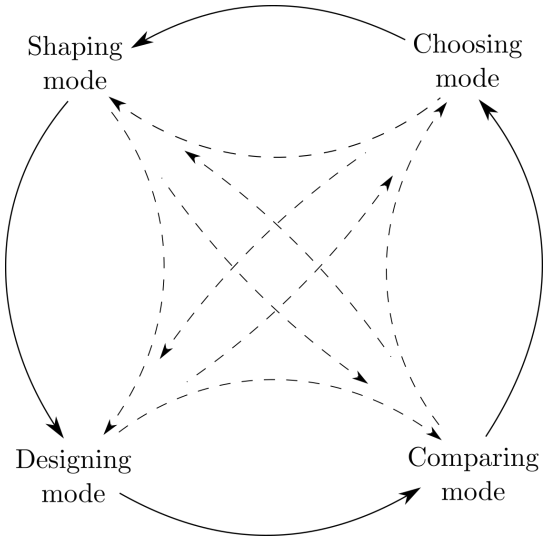


FIGURE 6.8: *A schematic representation of the SCA process, adapted from Friend* et al. *[178].*

Interestingly, the phases within which responses to these uncertainties are contained do not constitute a well-defined process of investigation. It is more accurately represented by an assembly of several logically successive activities. In other words, the SCA does not attempt to micro-manage the way in which its activities are implemented and conducted. The designers of the approach argue that the real world is too complex to justify a planning approach which incorporates an exact methodological process [177].

Strategic options development and analysis

SODA is a method dedicated to the methodic exploration of problems by means of cognitive mapping (*i.e.* by formally constructing a means-ends network). What differentiates it from all other cognitive mapping approaches is its use of *bipolar constructs* that facilitate the clarification of meaning. Similar to the SCA, SODA does not constitute a precise sequence of methodological steps. Instead, it is a collection of investigative techniques logically following one another. The SODA methodology contains four mutually interacting theoretical perspectives that inform these techniques, namely the *individual*, the *group*, the *consulting practice*, and *technology* [3]. The interaction among these perspectives is what drives the core notion of SODA — the construction of a representative, publicly viewable problem model, amenable to continuous change according to the development of the endeavour.

Problem structuring for multi-criteria decision analysis

In 2010, Stewart *et al.* [57] published an article titled *Problem structuring and multi-criteria decision analysis*. In it they presented the results of an attempt at connecting the methods of traditional operations research with the problems of the real world. Depicted in Figure 6.9, the proposed methodology was divided into five main categories, namely *identification of the problem*, *problem structuring*, *model building*, *use model to inform and challenge thinking*, and finally *develop an action plan*. This methodology will be built upon within the novel framework presented later in this dissertation.

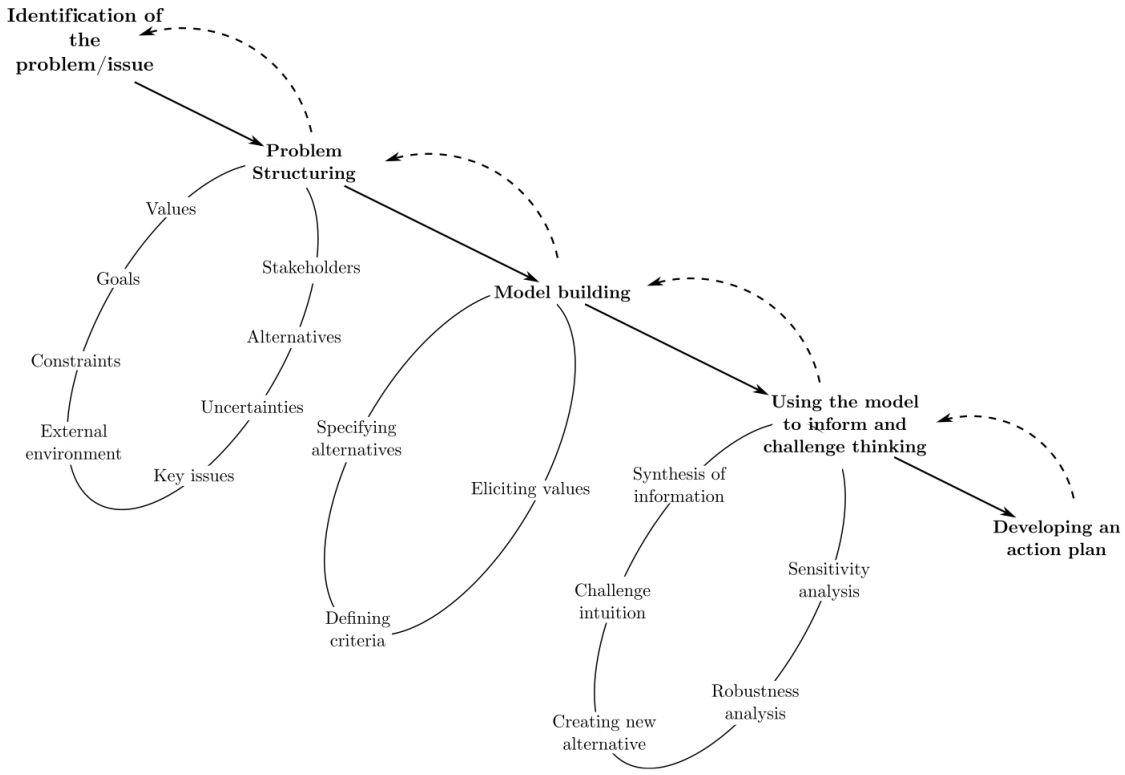


FIGURE 6.9: *A high-level overview of a soft operations research solution methodology [57].*

In conclusion, if cybernetics gave systems practitioners the capacity to tackle problems that are further along the dimension of non-human aspects of complexity, then soft systems thinking represented a step forward with respect to tackling problems that are further along on the dimension of complexities involving people (*e.g.* world-views, culture and interests). The scope of soft systems thinking is, however, not unlimited. A criticism levelled at soft systems thinking is the fact that it does not consider the effect of inequality within a given context. One can question the appropriateness of employing soft systems thinking in a context where several members hold the majority of power [234]. What then would be the point of discussion? Or should one rather ask why soft systems theories neglect to consider the influence of power, thereby neglecting the opportunity to guide analysts with respect to managing its influence. An interesting thought to note before closing is whether the existing paradigms of systems thinking may be combined in some sense to be more than the sum of its parts and supplement the respective limitations of the constituent paradigms.

6.7 Emancipatory systems thinking

For most of its history, the notion of emancipation was not one naturally associated with the objectives of management science. As Munro [333] explains, the concept of emancipation is one derived from the Latin word *emancipo*, which was employed by the Roman empire to refer to the liberation of a child from the legal control of its guardian. The publication of Werner Ulrich's *Critical heuristics of social planning* [458], however, stands as a landmark of the inclusion of emancipation within the scope of management science. More specifically, it marks the first time a systems approach recognised the need to pay attention to the influence of power upon the objectives pursued in a given context [234].

In pioneering this new paradigm, Ulrich distanced himself from the predominant theories of the time. He considered the existing theories limited in that they were only applicable when asking *how to do things*, never addressing the question of *what one ought to do* [234]. Toward addressing this latter question, the emancipatory approach embraces three major commitments, namely *critical awareness*, *social awareness*, and *human emancipation*. Critical awareness speaks to the importance of closely examining the assumptions and values that underlie any existing system or any newly proposed system design. Social awareness addresses the fact that analysts should be wary of the organisational and societal structures and norms that dictate the generally accepted manner of conduct. And finally, a commitment to human emancipation seeks to attain for all individuals the opportunity to progress in relation to their ability [232]. Within this general paradigm, Ulrich developed a practical methodology as a means to a more general end. To date, this methodology remains emancipatory systems thinking's greatest achievement [232] and so it is held to be representative of the approach and discussed accordingly.

Ulrich called his approach *critical systems heuristics* (CSH), interpreting each word in like manner to the well-known philosopher, Immanuel Kant. Being *critical* is to reflect upon the presuppositions that facilitate both the search for knowledge, and the pursuit of purposeful action. The acknowledgement of the *system* nature of the proposed methodology refers to the totality of elements (*e.g.* political, ethical, ideological and metaphysical) upon which decisions should be based. Finally, *heuristics* refer to the continual commitment to being critically aware of assumptions that govern activity [264].

CSH is built around the notion of what Ulrich called *boundary judgements* — those judgements that determine what is inside the system being studied and what belongs to its environment [233]. Essentially, these judgements elude to what the analyst believes to be relevant to the system under consideration. From the perspective of emancipation, these are key questions during the design of a system. In an attempt to address the lack of comprehensiveness that typically accompanied system designs, Ulrich formalised twelve questions designed to facilitate the process of boundary identification. The questions, listed in Table 6.3, are arranged around the distinction between those involved in the endeavour (*i.e.* the client, the decision maker and the designer) and those affected but not involved (*i.e.* the witnesses). The questions are designed to be asked in an *is* and an *ought* mode, contrasting the elicited responses so that the gap that exists between the way things are done and the way they ought to be done is illuminated [234].

Unlike soft systems thinking, CSH exhibits a sensitivity to the effect of inequality upon the process of problem formulation. Where soft systems thinking pays attention to the presence of conflicting interests [100], CSH goes a step further to recognise that the existing power distribution dynamically influences the manner in which conflicting interests are dealt with during an endeavour toward improvement. Essentially, CSH is a structured process of questioning the appropriateness of this effect of power upon the objectives pursued [234].

TABLE 6.3: *The boundary questions of the CSH approach, adapted from Jackson [232].*

Boundary question in *ought* mode	
1.	Who ought to be the *client* of the system?
2.	What ought to be the *purpose* of the system?
3.	What ought to be the *measures of success* of the system's performance?
4.	Who ought to be the *decision makers*?
5.	Which system *components* (*i.e.* resources and constraints) should be controlled by the decision maker?
6.	Which system *components* ought to be part of the systems environment (*i.e.* not controlled by the decision maker)?
7.	Who ought to be included as *designer* of the system?
8.	Which kinds of *expertise* ought to flow into the design of the system?
9.	Who ought to be be considered the *final voice* of the system's design?
10.	Who ought to belong to the *witnesses* representing the general citizens who will be affected by the system?
11.	To what degree and in what way ought the affected to be given the chance of *emancipation*?
12.	On which *world-views*, of either the affected or the involved, should the system's design be based?

Several critiques have been levelled against the CSH. The most dominant of these originated from philosophers and sociologists of a *materialist* persuasion. These critics argued that the CSH was built upon a philosophy that ignored the effect of the existing societal structures upon the ideologies that dominate societies at particular times [234]. In other words, CSH does not question the social structure of society, only the manner in which the existing structure is employed toward improvement. Because of this, it has been argued that the applicability of CSH is limited. It should be acknowledged, however, that these criticisms are rather harsh. Although CSH does not attempt to change the foundational structures of society, it goes some way to reveal when such structures are misused by those in power to benefit their own interests.

6.8 Critical systems thinking

Historically, *critical systems thinking* (CST) is a child of the 1970s, which is typically a period remembered for its musical genius, but fortunately, the genius of the late 20th century was not limited to the entertainment industry. Systems pioneers of the time are remembered for the radical shift in methodological approach they introduced during this period — a shift that came in the form of CST. Before the introduction of CST, the development of novel system methodologies had followed a particular pattern — critiques of existing approaches, followed by responses to these critiques in the form of newly developed methodologies [233]. CST was, however, a methodology of a different kind. The critique offered by the pioneers of the critical approach was not directed at the inherent value of a particular methodology, but was rather aimed at making salient the importance of matching methodology to problem. In other words, as a paradigm, CST recognised that solution methodologies are generally not *correct* or *incorrect*, but *well-suited* or *unsuited* to specific problem types (*i.e.* no single solution methodology is appropriate for solving every kind of problem). Thus, CST included within its intention the capacity to employ (when appropriate) other paradigms of inquiry [165].

This was a truly significant shift in the way practitioners had traditionally approached problem solving [469]. For the first time, the focus was not on the shortcomings of existing approaches,

but upon the collective value of employing methodologies when appropriate. As a means to that end, CST embraced five core commitments, namely *critical awareness*, *social awareness*, *human emancipation*, *wholesome development*, and *wholesome selection* [232]. Critical awareness represents the importance of closely examining the assumptions and values that are included within any existing system or any newly proposed system design. Similarly, social awareness addresses the fact that practitioners should remain weary of the organisational and societal norms and structures that dictate the generally accepted manner of doing. Human emancipation as a commitment, seeks to attain for all individuals the opportunity to progress in relation to their ability [164]. Wholesome development is the formal representation of the manner in which CST dedicates itself to incorporating alternative theoretical positions within its philosophy. Finally, wholesome selection represents CST's commitment to the complementary and informed use of alternative system methodologies.

As the reader might have noticed, the first three commitments listed above are identical to the commitments of emancipatory systems thinking (discussed in §6.7). This is because CST employed the core idea behind emancipatory systems thinking as inspiration and proceeded to develop that idea into the new paradigm typified by its fourth and fifth commitments. In 1991, two of CST's founding fathers, Flood and Jackson, published a methodology that became in every sense representative of the CST movement, namely a paper titled *Total systems intervention* [164]. It did not take long for total systems intervention to be heralded as a dominant approach when tackling planning and designing problems within the context of organisation theory. By 1995 the approach had received so much exposure to practice that Flood published an updated version of the methodology, this time under the new title *Local systems intervention* (LSI) [162]. In the remainder of this section, this updated version of the methodology is discussed.

A proper explication of the LSI methodology would be rather extensive. In order to shorten the discussion, attention is focused on the formally defined principles that embody the methodology and the actual process by which these are pursued, at the expense of discussing the philosophical foundations that constitute it (for a complete discussion, the reader is referred Flood [161]). There are four principles unique to LSI, namely *being systemic*, *achieving meaningful participation*, *being reflective*, and *working toward human liberty* [163]. Being systemic represents the belief that the world consists of interacting parts that function on several levels of abstraction. The second principle is rather self-explanatory — it represents a commitment to include all entities that are affected by the outcome of an endeavour within the endeavour. The reflective principle, on the other hand, describes two specific aspects that should be reflected upon during the endeavour: First, the nature of the relationship between divergent organisational interests that are pursued, and secondly, the reason for the dominant approaches being employed during the intervention. Finally, the principle of human liberty seeks for all the opportunity to progress according to their ability [233]. Interestingly, these principles interact and build upon one another. For example, the foundational principle of studying the world as if it were systemic is enabled and enriched by the second and third principles of meaningful participation and reflective inquiry.

As depicted in Figure 6.10, these principles are embodied by a process that involves a systemic cycle of inquiry, with iteration back and forth between three phases, namely *creativity*, *choice*, and *implementation* [165]. Starting with the creativity phase, the process in the clockwise direction works as follows. The main purpose of the creativity phase is to elicit issues to be dealt with by employing several methods of inquiry (*e.g.* brainstorming and descriptive organisational metaphors), which are then passed to the choice phase. The purpose of the choice phase is to select the methods that are best suited to manage the identified issues. Finally, the task of the

implementation phase is to connect the selected methods with the identified issues and to manage their interaction until a point in time at which specific change proposals have been identified. This outcome is once again passed to the creativity phase and the cycle is repeated [163]. In the anticlockwise direction, reflection on the appropriateness of the previously completed phase is formally investigated and established. Note that LSI can be thought of as a singularity — it is an integrated whole. Every phase focuses on one kind of activity, although it should be stressed that no phase exists independently [161]. The purpose of separating the phases is simply to facilitate a description thereof.

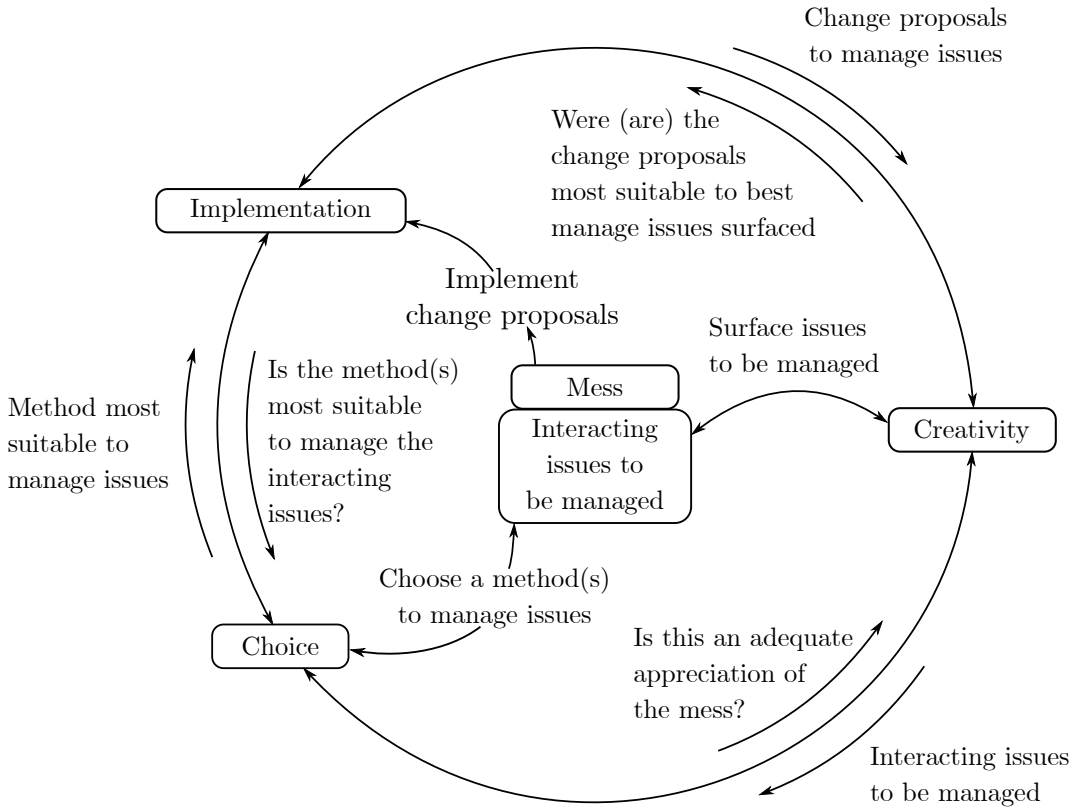


FIGURE 6.10: *A schematic representation of the LSI process [163].*

Note that the above description of the LSI methodology is an incomplete summary of its working. In its full glory, the methodology is rather extensive, providing guidance on how to uncover relevant issues, how to select appropriate methods related to those issues and even guidance on implementing the uncovered improvements. As mentioned before, however, an in-depth discussion on LSI is too broad in scope for the purpose of this chapter. The purpose of this chapter is to provide the reader with a basic understanding of the more general paradigms that exist within systems theory, not to describe the methods that function within these paradigms in detail.

In conclusion then, it is beneficial to describe the purpose for which the LSI methodology was designed. As Flood clarified in his formal critique of the initially proposed methodology (*i.e.* total systems intervention), this purpose is threefold: The design of effective and efficient processes and organisations, the debate of human and technical issues that arise in organisations, and the disempowerment of individuals from dominating organisational designs [163]. These objectives are essentially a response to three foundational questions of inquiry, respectively: *How should we do it*? *What should be done*? And finally, *why should it be done*? These questions constitute the core inspiration of LSI.

6.9 System dynamics

> *"Systems thinking is a discipline for seeing the structures that underlie complex situations . . . Ultimately, it simplifies life by helping us to see the deeper patterns lying beneath the events and the details" — Peter Senge*

In his book *Systems thinking: Creative holism for managers*, Jackson [233] selected the quote above to open his chapter on system dynamics, thereby exemplifying its ability to facilitate the discovery of those mechanisms that underlie the behaviour of a system. Within the systems movement, system dynamics has been accepted as a unique paradigm of endeavour. This is, however, somewhat surprising. Table 6.4 provides a summary of the key differences between the paradigms that have been discussed in this chapter. It can be seen that there is a significant difference in the manner in which each approaches problem solving. System dynamics, on the other hand, is better described as a unique modelling method, but in the view of the author there is no characteristic that justifies its inclusion as a unique paradigm within the systems movement. Within the context of this dissertation, systems dynamics is, therefore, taken to be an approach to modelling the non-linear behaviour of complex systems [234] and not a separate paradigm of endeavour. Accordingly, its working falls outside the scope of this chapter.

TABLE 6.4: *A summary of the key differences between the system paradigms described in this chapter.*

System paradigms	Key difference
Hard systems thinking	Traditional approach
Cybernetics	Black box modelling and feedback control
Machine learning	Autonomous learning
Soft systems thinking	Focus on the human component
Emancipatory systems thinking	Human liberation
Critical systems thinking	Combined paradigms described in this chapter

6.10 Chapter summary

There are several details that should be noted before closing the chapter. First and foremost, the manner in which the chapter has been structured may be misleading in respect of the relationships between the various system paradigms. For the sake of facilitating a description thereof, the system paradigms were chronicled in a categorical manner. The true state of things is, however, much more intertwined. From the origin of the systems movement, which was described in §6.1, its development has not been a perfectly sequential story. Its development can, in fact, accurately be likened to a fuzzy network of developments, all contributing to the overall development of the systems movement, as illustrated in §6.2. For example, the manner in which §6.2.1 was dedicated to the formal classification of methods into one of the paradigms reviewed was simply that — a formal classification of its core characteristics — many methods were developed within each paradigm, but have since evolved to the point where it is no longer possible to limit their working to a single paradigm.

Finally, CST was defined as a paradigm that endeavoured to employ every other paradigm in an attempt to match problem type to solution methodology. Similarly, the novel framework proposed later in this dissertation shares this objective. Consequently, it can be said that the work contributed in this dissertation belongs (for the most part) to the critical systems paradigm. Although the match is not quite perfect, it is one that is well suited to provide this work with a place to call home.

CHAPTER 7

Problem typologies

Contents

"What you perceive depends a great deal upon where you are standing" — CS Lewis

These words, penned by the British academic CS Lewis, require a certain measure of experiential understanding on the part of the reader. The truth of subjective perspective is one that is typically met with resistance by those who would rather live according to a limited, but absolute perception of the world — something system thinkers cannot afford to do. This chapter is dedicated to a review of the academic literature related to existing problem typologies which attempt to maintain a measure of objectivity in terms of solution methodology. A brief introduction in §7.1 sets the stage for the subsequent discussion on the specific problem typologies. In all, three problem typologies are discussed in §7.2, §7.3 and §7.4, respectively. Finally, the chapter comes to a close with a short summary in §7.5.

7.1 A typology of problems

Logic holds that the more complex a task, the more intricate the machine required to complete that task. Accordingly, the brilliance of the human mind suggests something about the context in which we live. The subjective experience of everyday life is exceptionally complex [87]. The general response to this complex flux called *life*, is not to simply experience it — we have an inherent desire to see life as being *meaningful* [270]. Akin to the economic status of an individual, *meaning* is a comparative measure — it can only be understood in comparison with something greater or less than itself. Consequently, a part of this meaning attribution is to conceptualise certain parts of our experiences as separate from other parts, to define certain boundaries that characterise a certain period of time as a *situation* [103]. These boundaries, which are the result of our subjective perception of reality, make it possible to assign a unique measure of meaning to separate periods of time. Furthermore, some of the situations we perceive move us to feel a

desire to change them, to improve them [103]. In order to improve these situations, it is often necessary to introduce some form of structure into the raw state of the world.

What is meant by *structure*? Certainly, more than will be discussed here. Nonetheless, when discussing such a multidimensional subject, the best place to start is at the beginning. Within the context of this dissertation, one of the initial stages of introducing structure into a problematic situation is believed to be the classification of problems into appropriate categories. The salient reason for distinguishing among different types of problems is the assumption that solving different kinds of problems requires distinct skill sets, that is, a unique solution methodology is associated with each type of problem. In this chapter, three of the more general problem typologies found in the academic literature are discussed.

7.2 Tame or wicked

In the year 1973, design theorists Rittel and Webber published an article that was to change the course of design theory [379]. In this article, titled *Dilemmas in a general theory of planning* [382], the authors argued that there existed a whole spectrum of problems that could not be solved using the traditional analytical solution methodologies of the time. It is important to note that at the time of its publication, traditional operations research techniques were being applied to every kind of problem. Unwittingly, the hammer and nail analogy, in which a problem is interpreted from the perspective of the solution methodology at hand, was the norm of the day. Rittel and Webber [382] were among the first to officially recognise the value of distinguishing between different kinds of problems. More specifically, they deemed it necessary to distinguish between two kinds of problems, namely *wicked* and *tame* problems.

At first glance, what Rittel and Webber meant by wicked problems is not apparent. This ambiguity is most likely the result of the multidimensionality of characteristics that define a wicked problem. Therefore, before attempting to define these characteristics it is beneficial to define the kind of problems that reside on the opposite end of the spectrum to wicked problems, namely tame problems [382]. There are several generally accepted characteristics associated with tame problems. These include a relatively well-defined problem statement, a definite goal, a solution that can be judged as being correct or incorrect, and finally, tame problems tend to remain solved once solved [379]. These characteristics pertain to problems of a Newtonian nature. The games of chess and go are tame problems, for example [208].

If a tame problem can be likened to the controlled and predictable flow of water from a tap, a wicked problem is appropriately likened to a mountain river after rain, it is nearly impossible to contain. Wicked problems are ill-defined, ambiguous and saturated with strong moral, professional and political issues [379]. Above all, wicked problems are constantly evolving — they are intricate, with interrelated issues that interact and evolve in a dynamic, social context [208]. In order to facilitate their identification, Rittel and Webber [382] defined ten criteria according to which wicked problems could be characterised. These criteria are discussed below.

7.2.1 No definitive formulation

For any given tame problem, it should be possible to provide an exhaustive formulation such that the information required to solve the problem is either explicitly stated or implicitly available [79]. When it comes to wicked problems, however, this is not the case. The problem statement of a wicked problem could never contain all the information required to solve it, as

the required information depends upon the adopted solution approach. Essentially, problem understanding and problem resolution are concomitant to one another [382].

Consider, for example, attempting to identify the root cause of poverty in third-world countries. It is tempting simply to say that poverty is caused by low income, and in part that is true, but one must proceed to ask what the determinants of low income are. Is it a deficiency in the cognitive or occupational abilities of the workforce, or is it a deficiency in the regional and national economies of the country? If it is the former, then the problem statement and subsequent solution should incorporate a country's educational system. Once again, however, within the education system one must investigate where the real problem lies. Does it include poor physical and mental health, spatial dislocation, cultural deprivation, or poor political and social skills? If an analyst were to succeed in formulating a problem such that he or she could capture the root causes of the difference between the present state of the world and what ought to be, then (s)he would thereby also have formulated the solution. In other words, what is being claimed is that the formulation of the wicked problem is to a large extent the problem itself [382].

Essentially, the process of formulating a wicked problem and developing its solution is equivalent, as every specification of the problem is fundamentally a stipulation of the scope within which solutions are considered. Because of this dependence between the definition of a wicked problem and the solution procedure, there is no definite formulation for a wicked problem [79].

7.2.2 No stopping rule

When two opponents face one another over a game of chess, they both have a very definite goal in mind, a very definite point at which the game will be either won or lost, namely the situation of checkmate. It is not so with wicked problems. The reason for this lies in the fact that solving a wicked problem is synonymous with understanding its nature [383]. The greater an analyst's understanding of the problem, the more effective solutions can typically be proposed. Unfortunately, there are no criteria for sufficient understanding, and no end to the casual connections that link interacting open systems, within a typical wicked problem space. Thus, an analyst can always do better — there is always the possibility that additional effort may lead to additional benefit.

In conclusion, the analyst pursuing a solution to a wicked problem does not terminate the search process once he has found a Pareto optimal solution. Instead, he or she terminates the solution process based on considerations external to the problem, such as running out of money, time, or patience, or stopping when the stakeholders are satisfied with a proposed solution. The final solution to a wicked problem is therefore given within the context of doing the best one can within the limitations of a project [382].

7.2.3 No theoretically correct solution

How does one qualify the correctness of a newly proposed fishing quota with respect to companies catching bluefin tuna in the pacific? Or the effectiveness of a newly developed plan toward reducing crime in the urban suburbs of a city? It is simply impossible to judge such proposals as being *correct* or *incorrect* [382]. Instead, the quality of such decisions is often only discovered long after their implementation [208]. When practitioners tackle wicked problems, they do not have the luxury of depending on a conventional criterion for objectivity in terms of which a solution can be deemed to be *correct* or *incorrect*. In fact, employing the term *correct* within

the context of wicked problems is of no value. To speak of a *correct* solution communicates a perfect fit between the problem and a proposed solution, it implies that the solution will cause the problem to have been solved indefinitely. Such solutions simply do not exist within the context of wicked problems.

Normally, when qualifying a solution to a wicked problem, there are several parties who are equally equipped and entitled to judge the quality of a solution, each prioritising conflicting goals. It is through the aggregate of these subjective opinions that the quality of a solution is typically determined. Consequently, it is more appropriate to use terms such as *good* or *satisfying* when describing the solution to a wicked problem [383].

7.2.4 No ultimate test of solution quality

When analysts investigate the quality of a proposed solution to a wicked problem, they do not consider the actual effect of implementing the solution — they rather consider the predicted effect of implementing the solution. This is because the impact of implementing a solution to a wicked problem can be likened to throwing a pebble into a pond, it will generate waves of consequences that are virtually unbounded [382]. Consider, for example, implementing a relaxed fishing quota with respect to catching bluefin tuna in the pacific. The full effect of such a change in strategy cannot be known at the time of implementation. It will likely take years before the local ecosystem will suggest a glimpse of its impact. Furthermore, the damage done when implementing such a solution can typically not be undone easily.

Essentially, the full impact of implementing a solution to a wicked problem cannot be evaluated before all the waves of repercussions have run their course [79]. And there is no practical way of tracing all the waves of influence to all the affected lives within a limited time frame. Hence, no ultimate test of solution quality exists.

7.2.5 No room for trial and error

Imagine a scientist tasked with the responsibility of developing a new kind of concrete. Within the development process it is the unique combination of theoretical understanding and the process of testing several prototypes that produces a final product. If the geometry of deformation of prototype 1 is not according to specification, the scientist can simply return to the drawing board and try again. Not so with the development of a solution to wicked a problem [383]. The solution to a wicked problem typically has irreversible, real-world consequences — trial-and-error is simply not ethically defensible [79]. One cannot build a freeway in order to determine its effectiveness — such an action leaves traces that cannot easily be undone. Whenever the consequences of implementing a solution have an irreversible influence on the context of the real world and the lives of people within that context, chances are that one is dealing with a wicked problem [382].

7.2.6 No exhaustively describable set of solutions

The element of surprise that exists within a game of chess is not the product of a player executing a move that has never been attempted before — instead it is simply the failure of an opponent to anticipate such a move. This is because the game of chess has a limited set of rules within which every possible situation that can occur is contained. In a technical sense, the word *innovative* has no place in the context of chess. Not so in the world of wicked problems. An exhaustive set of strategies for dealing with crime in the streets, for example, never has, and never will be

documented [382]. The life of social policy does not function according to such a set of definable rules, or rather if it does, humanity has failed in the pursuit of its identification. In such fields of ill-defined problems and consequently ill-definable solutions, the set of feasible solutions depends upon the capacity to judge new ideas realistically [383].

7.2.7 Every wicked problem is essentially unique

No matter how similar, for any two problems at least one distinguishing property should be identifiable, for otherwise the two problems are, in fact, the same problem. Every problem, therefore, is unique in a trivial sense, although something quite different is meant by *essentially unique* [79]. As Rittel and Webber [382] put it, no matter how long the list of similarities between two wicked problems, there will always be an additional distinguishing characteristic that is of overriding importance. This quality of essential uniqueness is part of what makes solving wicked problems a form of art — no amount of experience will ever render a wicked problem simple to solve, a measure of innovation will always be required.

To put it another way, there are no classes of wicked problems in the sense that the salient principles of a solution can be adapted to fit all members of that class. For example, the conditions of a subway construction project in two cities may look similar, but the designers would be ill-advised to apply the same blueprint to both cities. Differences in commuter habits could far outweigh the apparent similarities in the physical conditions surrounding the project. Essentially, despite the apparent similarities among wicked problems, one can never be quite certain that the particulars of the problem do not override the commonalities it shares with problems encountered before [208].

7.2.8 Wicked problems are consequential

If one finds oneself on one side of a river while actually wanting to be on the other, without the means of getting there, one has a problem. If one cannot find a polite way to formulate the fact that there is no potential in a proposed business plan, one has a problem. Whenever there is a gap between the current state of affairs and the state as it ought to be, and there is no salient way of bridging that gap, one has a problem [207]. The process of solving a problem starts with the business of identifying a casual explanation of the discrepancy (*i.e.* the gap between where one wants to be and where one is), followed by an attempt to eliminate this discrepancy [383]. In the context of wicked problems, however, this process typically elucidates a new problem of which the original problem is a symptom. Again, this new problem can be considered a symptom of yet another, higher-level problem. That is the nature of wicked problems — they can each be considered to be the symptom of another problem [382]. Crime in the streets, for example, can be considered to be a symptom of several higher-level phenomena, such as moral decay, the distribution of wealth, or whatever casual explanation one considers defensible.

This multi-level quality of wicked problems implies that there is not a *natural* level on which to solve wicked problems. During the solution process, it is simply advisable that an analyst maintains the balance between addressing critical symptoms, and making the problem too general, and thus, too complex.

7.2.9 The analyst has no right to be wrong

In *The logic of scientific discovery*, Karl Popper [365] proposed as scientific principle the fact that solutions to problems are only hypotheses offered for refutation. This proposal was grounded in

the fact that, by definition, there are no proofs for hypotheses, only potential refutations. The longer a hypothesis can withstand the test of time, the better its corroboration is considered to be. Accordingly, the scientific community typically does not hold academics accountable for proposed hypotheses that are later refuted. In the context of wicked problems, no such immunity is morally defensible [382].

When solving a wicked problem, the purpose is not to find truth. It is rather to narrow the gap between the present state of the world and the way it ought to be — to improve some characteristic of the world in which people live. Accordingly, analysts are accountable for the consequences of implemented solutions.

Rittel and Webber [382] proposed one further defining characteristic of a wicked problem: "The existence of a discrepancy representing a wicked problem can be explained in numerous ways. The choice of explanation determines the nature of the problem's resolution." This characteristic has, however, been deemed redundant as is fails to narrow down the scope of wicked problems in any additional way and is therefore not discussed here.

Before concluding this section, however, consider the effectiveness of the typology of Rittel and Webber for classifying problems in a manner that facilitates a solution process. At the time of its formulation, the purpose of this typology was to shed light upon the fact that not all problems could be considered equal in nature — that the same solution methodology could not be applied to every problem encountered. For this purpose, the typology is perfectly suited, but for the purpose of elucidating the more nuanced elements influencing a problematic situation, the typology is of limited use. Nonetheless, it must be acknowledged that the typology provides a valuable foundation for the further classification of problems.

7.3 The dimensions of a problematic situation

In the world of design, whether it be the design of a new product or a novel problem typology, it is often the synthesis of many, different ideas from several sources that yields the most ingenious results. Following this logic, the second typology discussed here was synthesised more than it was formulated. In his PhD dissertation titled *Student perceptions of problems' structuredness, complexity, situatedness, and information richness and their effects on problem-solving performance*, Lee [282] described the manner in which he noticed an ever-increasing agreement in the academic literature in respect of the dimensions that define a problem. Drawing on this academic consensus, Lee proposed a framework that, he argued, included the core elements characterising a problem, namely *structuredness*, *complexity* and *situatedness*. Unlike Rittel and Webber [382], Lee did not coin these terms, nor did he define them in a novel way. Instead, Lee endeavoured to synthesise the work of several academics into a single framework. More specifically, the work of Wood [490], Jonassen [250], Meacham *et al.* [312] and Bassok [49] formed the foundation of Lee's proposed problem typology. In the remainder of this section the dimensions of a problem, as synthesised by Lee [282], are discussed.

7.3.1 Problem Structuredness

Lee [282] elegantly described a problem's structuredness as the intrinsic value of its internal structure. Arguing that this internal structure found form through a unique blend of several problem elements, including its *initial state*, *goal state*, *constraints* and *operators*, Lee claimed that the degree of a problem's structuredness can be determined by recognising the nature of these elements [15, 248, 490]. The history of determining a problem's structuredness is an

interesting one — it tells the tale of several academics working toward its formal coinage. It was Wood, however, who in 1983 was the first to formally define this dimension in an article titled *Inquiring systems and problem structure: Implications for cognitive development* [490].

Wood established the spectrum of structuredness by defining the extremes that contain it. On the one end of the spectrum, he coined the term *well-structured* and on the other end the term *ill-structured*. Following this terminology, several academics have defined and redefined these categories. Lee [282], however, adopted the following definition. A well-structured problem is one that presents all the information required for its solution in the problem representation, contains a known goal state, requires a limited number of regular and bounded rules which are organised in a prescriptive and predictive manner, possesses a correct, convergent answer and admits a prescribed solution process [250, 490]. An ill-structured problem, on the other hand, is one that typically exhibits a highly intricate structure and cannot be solved in predictable or obvious ways, requires solutions that are interdisciplinary in nature, involves conflicting objectives, and contains a considerable amount of uncertainty [247, 312].

When defined in such a formal context, the difference between well-structured and ill-structured problems may seem apparent, and yet the intricacies of their real-world contexts often cause decision makers to hesitate when characterising the structuredness of a problem. Accordingly, the process of elucidating a problem's structuredness should be approached with caution.

7.3.2 Problem complexity

The stem of the word complexity (*i.e.* complex) is composed of the Latin words *com*, meaning "together" and *plex*, meaning "woven" [314]. In general, therefore, complexity is characterised by the degree to which the respective elements of a system are interwoven and interrelated [282]. In order to incorporate this dimension into the proposed problem typology, Lee had to make its definition specific to the context of real-world problems. To that end, Lee [282], once again, drew upon and combined the work of several authors. Among these, Lee paid special attention to definitions proposed by Wood [491] and Jonassen [251], which are discussed below.

Wood [491] defined complexity along three dimensions: *component complexity*, *coordinative complexity* and *dynamic complexity*. Respectively, these dimensions represent the number of distinctly recognisable acts required to perform a task, the nature of relationships among these tasks, and the degree to which the relationships among these tasks are transformed by changes in the problem's external environment. In other words, Wood believed that complexity is a function of every problem component, the reciprocal interactions between these components and the effect of external factors on their interactions.

Jonassen and Hung [251] built upon Wood's [491] initial definition and described problem complexity in terms of what he called relevant internal and external factors. Internal factors are those on the problem solver's level of experience as well as the breadth and depth of domain-specific knowledge required to solve the problem. External factors, on the other hand, include the relational complexity among domain concepts, the problem's level of transparency (*i.e.* the number of unknowns in the problem space), the heterogeneity of possible problem interpretations, and the legitimacy of alternative solutions.

Interestingly, Lee did not select a single definition, but employed a synthesis of the definitions described above in a casual manner. Built upon these definitions, Lee incorporated a classification proposed by Rose [476], which categorises the continuous spectrum of complexity into one of two categories, namely *simple* or *complex*. These terms find meaning in view of the definitions described above. To illustrate this point, consider the definition proposed by Wood [491],

who defined complexity along three dimensions, namely component, coordinative and dynamic complexity. A complex problem will be characterised by a high rating with respect to each of these dimensions.

There is an interesting thread of discussion in the academic literature surrounding the overlap between the complexity and the structuredness of a problem [248]. Intuitively, it follows that the less apparent the inherent structure of a problem context, the more intricate one can expect the relationship between elements to be. A low level of structuredness does not imply no structure. Instead, it implies that the connections between these dimensions are exceptionally difficult to elucidate in the sense that structuredness represents the network of dimensions in the problem space. In other words, structuredness describes the ease with which the elements of a problem and the casual relationships between these elements are identifiable, while complexity describes not only the depth of relationships between elements, but also the nature of those relationships and the influence of external elements on those relationships. Clearly, although these dimensions overlap, they are rightfully defined as two separate dimensions in the problem typology proposed by Lee [282].

7.3.3 Problem situatedness

As Wood [490] so elegantly put it, "The context in which a problem is embedded becomes a significant part of the problem and necessarily part of its solution." Evidently, it is the influence of context that defines the situatedness of a problem. In other words, the situatedness of a problem represents the meaningfulness of a problem's context as interpreted by a problem solver's prior knowledge and real-time experience while solving a problem [13, 387]. Intuitively, it follows that the context of a problem (*i.e.* the situatedness) is related to problem solving performance [249]. The academic literature is, however, rather ambiguous in attempts at describing this relationship. For example, Chi *et al.* [107] argued that the more situated a problem turns out to be, the easier it is to define and represent it, and therefore, results in an increased possibility of successful problem solving. In contrast, Jonassen [250] argued that workplace engineering problems are made more difficult to solve because high levels of situatedness result in unanticipated problems — a situation may be constrained by several overlapping contexts.

Lee [282] did not, however, consider an exact definition of the relationship between situatedness and problem solving performance critical. His priority was simply to accommodate this dynamic relationship in the proposed typology. Accordingly, Lee adopted the following categorisation of situatedness: The *contextualised problem*, which is highly dependent on the dynamics and flexibility of the external environment, and a *decontextualised problem*, which lacks meaningful contextuality within itself [28, 297]. Similar to structuredness and complexity, these categories do not constitute a dichotomy, but represent the extremes along a spectrum of situatedness.

The decontextualised problem is the classic, theoretical problem typically found in academia. It does not require domain-specific knowledge for solving, nor is there a correct moment in relation to the external environment when decisions toward problem solving should be made. The elements (*i.e.* the initial state, goal state, constraints, and operators) are not situated, they are abstract and isolated. And finally, the solutions to decontextualised situation problems are generally accepted. The contextualised problem, on the other hand, takes place in the real world, has an underlying story and emerges from a specific domain. Problem solvers are expected to respond to the problem in a time-appropriate manner — decisions should be made at the appropriate moment with respect to the needs of the external environment. Finally, the solutions to highly situated problems remain fuzzy, coated in a measure of uncertainty [13, 312, 490].

Consider the effectiveness of this dimension within Lee's [282] proposed typology. Interestingly, the effectiveness of situatedness depends, by definition, upon the defined boundaries of a system. The situatedness of a problem varies as the boundaries which define the problem varies. Furthermore, it seems as if the definition of situatedness could be included within the definition of complexity, the only difference being the location of application. For example, the situatedness of a system can be redefined as the nature of interaction between the elements within the system that defines the problem, and the elements outside the system, which is rather similar to the definition of complexity. In other words, it would seem that this dimension may be redundant, not because of the characteristics it attempts to categorise, but because the scope of problem complexity includes the capacity to categorise this characteristic. Lee, however, evidently thought differently, considering situatedness worth including within his typology [282].

7.3.4 Problem information richness

In 1991, Confrey [118] described problem solving as the act of taking the information describing a problem and synthesising it in a novel way. If this claim is accepted, it follows that there exists a dynamic relationship between problem solving performance and the information richness of a problem. Intuitively, this seems to make sense: A lack of information may certainly be the source of ambiguity with respect to the nature of problem elements. Initially, Lee agreed with this logic and considered information richness worth including in his proposed problem typology, classifying problems as either *information lean* or *information rich*, as defined by Quesada *et al.* [371].

After the validation phase of his research, however, Lee [282] concluded that information richness did not add enough latitude to the typology to justify its inclusion. Hence, it was not included in the final typology. This might come as a surprise — intuition informs our logic of the value of information toward solving problems. It should be noted, however, that Lee did not claim that information has no value, but instead that the first three dimensions defined above, adequately describe the information richness of a problem.

In conclusion, Lee's [282] final typology incorporated three dimensions, namely structuredness, complexity and situatedness. According to the discretely defined categories within each of the dimensions, the typology contains eight categories into which problems can be classified. While these categories do well in describing the nature of the system within which a problem is embedded, it falls short of adequately describing the nature of the social aspects of a problem. For the purposes of this dissertation, an adequate representation of the social reality of a problem is critical to successful problem solving. Consequently, the next typology discussed is one that pays more attention to the *soft* side of problematic situations.

7.4 Beyond a system of systems methodologies

The final typology reviewed in this chapter, proposed by Jackson and Keys [229], was developed over many years (1984–1990), with several versions being published during this time. This literature study is, however, only concerned with the final version of the typology as published in the 1990 article titled *Beyond a system of systems methodology*.

The inspiration for the proposed problem typology can be traced back to the work of the renowned academic, Russel Ackoff. More specifically, it can be traced back to a list of complications that often occur within problematic situations as documented by Ackoff [5] in his well-known work titled *Scientific method: Optimising applied research decisions*. By examining

this list, Jackson and Keys determined that complications to problems (*i.e.* issues that make problems tricky to solve) are a consequence of changes in either the nature of the system in which a problem is located, or the nature of the decision makers participating in the solution process [231]. Consequently, Jackson and Keys determined that these two dimensions were an effective way of classifying problems.

7.4.1 Complexity

Over the years, academics have categorised the nature of the system within which problems are embedded along several dimensions. Notably, *complexity* has found itself employed within these classifications more often than not. Jackson and Keys followed this school of thought and employed a *mechanical-systemic* classification to describe the level of complexity associated with the system that provides context to a problem. After selecting these categories, Jackson and Keys issued an interesting note of caution — depending on the point of view adopted, the same system can be viewed as mechanical or systemic (*i.e.* simple or complex) [229]. Take the labour market, for example, which when contemplated from a highly aggregate point of view, can be treated as a simple supply-demand system. This highly aggregated point of view is typically adopted when the problems being studied are related to macro-economic issues like inflation, levels of export trade, or interest rate fluctuations. But the labour market can also be studied at a much lower level of abstraction, with individuals acting as parts of a highly complex system. This micro point of view is typically adopted when the problem being addressed concerns the individual [117].

Granted this observer-dependent nature of the mechanical-systemic criterion, Jackson and Keys provided a brief description of the main characteristics that characterise the complexity of a problem. On the one end of the spectrum, a mechanical system is typically perceived as consisting of a small number of elements that interact with one another in regular and predictable ways [229]. Within the context of a systemic system, however, the attributes that constitute its parts and the relationship between these parts are not directly observable. As a result, it is difficult to understand the nature of the system (*e.g.* the network of the problem causations or the possible effects of implemented solutions). Furthermore, Jackson and Keys make it clear that a systemic system will change as a function of time, the evolution of which can typically be accredited to the continuous interaction between a systemic system and its environment (which is continuously changing as a function of time). As a result, if a representation of the law governing the interaction between the parts of the system can be derived, it will always be probabilistic in nature. Finally, Jackson and Keys concluded their description of the prototypical systemic system by stating that such a system will be influenced by behavioural factors (*e.g.* political, cultural or ethical issues) [229].

7.4.2 The nature of progress

A classification of problem contexts would be incomplete without addressing the human component that saturates real-world problems. People are exceedingly intricate beings with countless dynamic, internal workings. Fortunately, many of these dynamics are not detrimental to successful problem solving. Accordingly, the right set of classifying dimensions should elucidate the dynamics that are pertinent to the process of problem solving, while ignoring the dynamics that can be considered irrelevant [229].

Within the context of Jackson and Keys' work, the nature of progress is considered to be one such dynamic [231]. More specifically, the question of interest is concerned with the tactics required for the formal establishment of the objectives to be pursued within an intervention. Consider,

for example, a problematic situation where participants are in genuine agreement with respect to the objectives to be pursued, have compatible values and beliefs, and share common interests. It follows that the tactics to be employed to encourage progress within such a context will be very different from a context in which, for example, participants have conflicting objectives. Accordingly, where there exists genuine agreement among participants, a problematic context can be described as being *unitary*. If the state of things can be described by conflicting world-views and, to some extent, divergent objectives, but it remains possible to facilitate discussion to a point of genuine compromise and agreement, a problematic context can be described as being *pluralist*. Finally, if there is limited common interest among participants, if there is a fundamental air of conflict, and the only way in which progress can be made is through the exercise of power and domination, a problematic context can be called *coercive* [232].

It is worth mentioning that the final aforementioned category within the nature of progress (*i.e.* coercive) was not present within the initial formulation of the proposed typology. It was only in the years following its first publication that practitioners provided the necessary feedback to justify its inclusion. Practitioners found that identifying the coercive tactics employed within a problem space is sometimes critical to selecting an appropriate solution methodology [231]. Jackson and Keys responded to this development by adding a third and final category to this dimension in order to classify the coercive character of problematic situations.

Thinking of problematic situations, it is intuitively true that most, if not all, such instances are accompanied by a measure of competition. The manner in which participants elect to respond to this competition forms the foundation of subsequent interaction. It is, therefore, fitting that Jackson and Keys incorporated the nature of progress into the proposed problem typology, something that was rather novel in the world of problem typologies.

At the time of its formulation, the typology proposed by Jackson and Keys was rather innovative on several accounts. Most notably, the attention it granted the human component of professional problem solving was rather uncommon at the time. This typology was published in the era following the 1950s and the *dreary sixties*, a period during which traditional systems thinking had retreated from the foreground of social issue management [232]. This makes their pioneering effort back to the issues of strategic social management all the more commendable. Finally, this typology is deemed rather mature in its mission to accommodate the identification of issues that, if ignored, may potentially undermine an employed solution methodology.

7.5 Chapter summary

The manner in which our understanding of the world is framed, determines the effectiveness of our response to its mysteries. In order to respond appropriately to real-world problems it is, therefore, of paramount importance to frame our understanding of such problems appropriately. In order to simplify this task, academics have endeavoured to formulate typologies which effectively elicit the most influential characteristics within a problem space. After being introduced in general in §7.1, three such typologies were reviewed, namely *tame or wicked* in §7.2, *the dimensions of a problematic situation* in §7.3, and *beyond a system of systems methodologies* in §7.4. These typologies were evaluated in terms of their effectiveness and it was recognised that the efficacy of a typology will, to some extent, always depend upon the appropriateness of its use within a specific context. No single typology has the ability to elicit all the relevant characteristics of a problem space for every problem type. Practitioners would do well to remember that any typology remains but a tool in the practitioner's toolbox, and so it remains the responsibility of the practitioner to utilise it appropriately. Nonetheless, the typologies discussed in this chapter are of high quality.

CHAPTER 8

Methods of stakeholder selection theory

Contents

It is often difficult to distinguish the relevant from the irrelevant in the context of real-world problem solving. Frequently, more people are affected by the occurrence of a problem than can practically be included within the process of its resolution. Simply put, the act of real-world problem solving involves the selection of stakeholders — a process that has proven critical to the sustained resolution of real-world problems [172].

In an article titled *Who's in and why? A typology of stakeholder methods for natural resource management*, Reed *et al.* [376] argue that the methods of stakeholder selection theory may be classified into one of three categories according to whether they are designed to aid analysts in (1) *identifying stakeholders*, (2) *differentiating between and categorising stakeholders*, or (3) *investigating the network of relationships among stakeholders*. In the years since its formulation, many theorists have adopted the above classification and, as such, it is employed to organise the work presented in this chapter. Accordingly, §8.1 is dedicated to an analysis of stakeholder methods that belong to the first category listed above, while §8.2–§8.5 are dedicated to an analysis of methods that belong predominantly to the second category listed above, and §8.6 is dedicated to a discussion of stakeholder methods that belong to final category of methods listed above. In addition, §8.7 contains a high-level comparison of the stakeholder methods discussed according to the qualities that distinguish them. Finally, the chapter is brought to a close with a brief summary in §8.8.

8.1 Focus groups, interviews and methods of sampling

In the article mentioned above, Reed *et al.* [376] argue that the methods contained in the stakeholder selection literature fail to address the difficulty of stakeholder identification adequately —

that methods designed to facilitate the selection of stakeholders focus predominantly on either the categorisation of pre-identified stakeholders, or the analysis of the network of relationships among them, without ever addressing how analysts are to go about identifying the stakeholders involved in such an analysis [376]. As this section is dedicated to the description of methods that belong to this, less well-addressed activity of stakeholder identification, there are a limited number of methods to review and, as such, it is possible to present the entire collection of methods employed by analysts to facilitate the activity of stakeholder identification. More specifically, by synthesising the methods proposed by Reed *et al.* [376] in 2009, and by Valerio *et al.* [462] in 2016, to be representative of the academic literature related to stakeholder identification, the following methods of stakeholder selection are considered: *Focus groups*, *semi-structured interviews*, *snowball sampling*, *purposive sampling* and *convenience sampling*. Although this list of methods may seem substantial, it will become clear during the exposition that, as Reed *et al.* suggest, they often fail to address the issue of stakeholder identification adequately. Before commencing their description, it is brought to the reader's attention that the aforementioned methods are only discussed to the extent that their ability to facilitate the identification of stakeholders is made salient.

The description of the term *focus group*, proposed by Freeman in an article titled *Best practice in focus group research: Making sense of different views* [173], suggests that a focus group is best characterised as a kind of group interview with strategically selected individuals according to "a carefully designed topic guide," with the purpose of eliciting information in respect of a specific issue of interest [173]. This description of a focus group is consequential as it reveals the dependence thereof upon secondary methods of investigation (*i.e.* "a carefully designed topic guide"). In other words, when stakeholder theorists suggest the use of a focus group as a method capable of independently facilitating the identification of stakeholders, they misinterpret the nature and purpose thereof. More specifically, a focus group is more accurately described as a general approach to qualitative enquiry and not a methodology designed specifically for the identification of stakeholders. The lack of a "carefully designed topic guide" to facilitate stakeholder identification thus remains an unaddressed issue despite the notion of a focus group. Accordingly, the discussion progresses to the next method of interest.

In a book titled *Key methods in geography*, Longhurst [296] states that a semi-structured interview represents a kind of guided verbal interchange between an interviewer and an interviewee with the aim of gaining insight into a certain topic of interest [296]. The idea behind this kind of interview seems to revolve around the construction of a list of predetermined questions designed to guide an interview in such a way that interviewers are able to respond and adapt to a discussion as it unfolds [136]. Unfortunately, despite its potential in other spheres, based upon the above description, this method of analysis does little more to address the issue of stakeholder identification than the one described before. To elucidate, the list of guiding questions mentioned above may be interpreted to represent a special kind of methodology according to which analysts attempt to structure their enquiry. The construction of these questions is, however, left to an analyst and, as such, in like manner to a focus group, a semi-structured interview is better described as a general approach to the acquisition of knowledge, and not a methodology specifically designed to facilitate the identification of stakeholders [136].

Although elementary, the next approach to stakeholder identification, called snowball sampling, is certainly more of a self-contained method for the identification of stakeholders than those described before it. In essence, snowball sampling is a non-random sampling technique in which primary data sources (*i.e.* an initial group of key stakeholders), identified according to a set of distinguishing characteristics, are employed as *tracers* to identify additional data sources by means of *chain referral* [64]. More specifically, the initially identified key stakeholders are asked

to peruse the network of their connections in an attempt to identify entities that conform to a set of distinguishing characteristics deemed descriptive of stakeholders [75]. Despite being a method specifically designed to facilitate the form of sampling that is stakeholder identification, snowball sampling leaves much unaddressed. For one, the criteria according to which analysts should identify relevant stakeholders is left open to interpretation and, in addition, no aid is given as to how entities are to go about exploring the network of their connections. Because of these shortcomings Reed *et al.* [376] presumably considered the task of stakeholder identification inadequately addressed despite the existence of the method of snowball sampling.

The final two sampling methods, called purposive sampling and convenience sampling, respectively, are rather uninvolved. The first of these "methods" simply suggests that analysts rely upon their subjective judgement when distinguishing relevant from irrelevant stakeholders [147] and, in an equally one-dimensional fashion, the second suggests that analysts should select stakeholders according to the convenience of their availability, and nothing else [155]. Accordingly, Reed *et al.* [376] warns that the use of either of these methods may result in an excessively homogeneous selection of stakeholders (adding that the success of their application should improve as analysts become more experienced in their use) [449]. If anything, the methods of purposive sampling and convenience sampling seem to be well-suited to the kind of endeavour in which stakeholder identification is both a trivial and inconsequential task. Unfortunately, however, stakeholder identification is often not only exceedingly difficult, but consequential in nature [242] and, as such, it remains a task inadequately addressed by the methods of stakeholder selection theory. For whatever reason, there is a clear gap in the academic literature in this regard — a gap that the work presented later in this dissertation attempts to fill.

8.2 Toward a theory for stakeholder identification and salience

In the years that followed the release of Freeman's oft-cited book titled *Strategic management: A stakeholder approach* (1984) [172], the notion of a stakeholder became rather popular among those working in the sphere of corporate management [325]. Unfortunately, however, in addition to the unaddressed difficulty of practically identifying stakeholders, as discussed in §8.1, the popularity of the term did not coincide with a general understanding of precisely how analysts could go about determining the relative relevance of two dissimilar stakeholders. It was only in 1997 (15 years later) when Mitchell *et al.* [325] proposed a framework for the classification of potential stakeholders toward facilitating their case-specific selection, that significant progress was made in that regard. The core contribution made by Mitchell *et al.* [325] was twofold: First, they proposed three dimensions according to which analysts could attempt to determine the relevance of an entity as a stakeholder, namely *power*, *legitimacy* and *urgency*, and secondly, they formalised a typology of unique stakeholder classes according to which they proceeded to prescribe which kind of entities managers should consider as stakeholders[1]. In the remainder of this section, these contributions are discussed in turn, starting with the first dimension defined to govern the notion of stakeholder relevance (*i.e.* power).

Since the mid-20th century, most definitions of power have been based upon the early Weberian notion that the measure of an entity's power essentially represents the probability that the entity in question will be able to carry out its will in the face of an adversary [477]. In general, it would seem that Mitchell *et al.* [325] adopted a similar concept of power, but in order to describe their notion of power in a manner that actually facilitated the consideration of a entity's relevance as a stakeholder, the authors had to endeavour beyond such a general definition and ask a more

[1] The stakeholder classes proposed by Mitchell *et al.* [325] were based on the aforementioned dimensions.

fundamental question: What grants an entity the actual ability to exert its will, or alternatively, what are the actual bases of power? To answer this question, Mitchell *et al.* turned to the work of sociologist Amitai Etzioni who, in 1964, defined the notion of power along three dimensions, namely *coercive* power, representing power based on the physical resources of force and violence, *utilitarian* power, representing power based on material or financial resources, and *normative* power, representing power based on symbolic resources. In other words, Mitchell *et al.* argued that the measure of an entity's power can be determined according to the extent to which it has, or has access to, coercive, utilitarian and normative means by which to impose its will (see Etzioni [148] for an in-depth discussion).

The second dimension of stakeholder relevance proposed by Mitchell *et al.* [325] (*i.e.* legitimacy) is rather less precisely defined. In essence, Mitchell *et al.* determined that the notion of power differs from that of *authority* in that authoritative entities are those that, in addition to being powerful, are at liberty to employ their power according to the social norms that underlie their context [477]. By interpreting this dynamic, the authors argued that an entity with authority is a more relevant stakeholder than merely the powerful and, as such, defined the notion of legitimacy to represent the extent to which the consideration of an entity's priorities, and thus its inclusion as a stakeholder, is in line with the socially accepted norms of the prevailing context (*i.e.* the socially constructed system of beliefs, values and definitions to which the instance of stakeholder identification relates) [438]. In addition to an entity's power and legitimacy, Mitchell *et al.* proposed one final dimension of stakeholder relevance, namely urgency. Simply put, the notion of urgency is intended to facilitate the classification of stakeholders according to how "pressing" their demands are. More specifically, the urgency of an entity's demands are determined according to (1) how time-sensitive they are, and (2) how important their realisation is (as subjectively determined by the related entity) [325].

By synthesising the dimensions described above, Mitchell *et al.* [325] put forward a typology of stakeholder classes containing seven unique categories, namely a *dormant stakeholder*, a *discretionary stakeholder*, a *demanding stakeholder*, a *dominant stakeholder*, a *dangerous stakeholder*, a *dependent stakeholder* and a *definitive stakeholder*[2], as depicted in Figure 8.1. Based on these classes, Mitchell *et al.* proceeded to argue for, or against, the consideration of entities as stakeholders. In general, they did so according to the following proposition: Stakeholder salience will be low, moderate or high depending on whether managers perceive one, two, or three of the aforementioned stakeholder attributes — power, legitimacy or urgency — to relate to a specific entity, respectively [325]. For a more in-depth discussion on the arguments made in respect of each stakeholder-class, consider the following summarised version thereof.

As may be seen in Figure 8.1, dormant stakeholders are characterised by a possession of power, but the lack of both legitimacy and urgency. According to Mitchell *et al.*, such entities remain inactive (*i.e.* dormant) with respect to a related firm's activities and, as such, should not be considered as stakeholders to the firm. As a concluding word of caution, however, the authors do advise managers to keep such entities in their periphery, in case they acquire additional stakeholder attributes. The second group of stakeholders, called discretionary stakeholders, possess the attribute of legitimacy, but neither power nor urgency, which makes them a rather interesting class of stakeholders from a social responsibility point of view (*i.e.* they are most likely to be recipients of what Caroll [91]) calls *corporate philanthropy* [488]. Regardless, the key point made by Mitchell *et al.* in respect of discretionary stakeholders is that, in the absence of power and urgency, there is "absolutely no pressure on managers to engage in an active relationship

[2]To clarify, note that Mitchell *et al.* employ the term stakeholder to describe dissimilar kinds of entities, regardless of whether or not they actually argue for the consideration of said entity as a stakeholder to a firm (in a practical sense) [325].

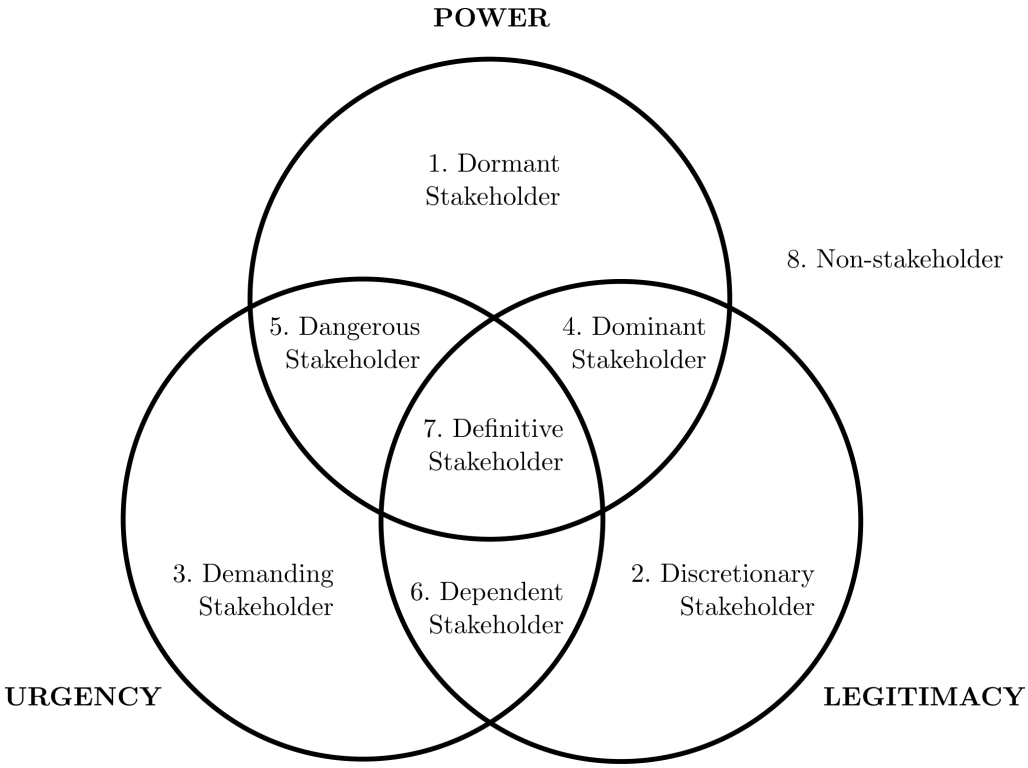


FIGURE 8.1: *A classification of unique stakeholder classes according to the dimensions of power, legitimacy and urgency [325].*

with such a stakeholder" [325]. Interestingly, Mitchell *et al.* do not pay much attention to the class of stakeholders characterised solely by the urgency of their claim, called demanding stakeholders, describing them to be the mosquitoes buzzing in the ears of managers: Irksome, but not dangerous, bothersome but not warranting significant attention, if any at all [325]. To conclude their discussion of such stakeholders, the authors describe a lone millenarian picketer marching in front of a firm's headquarters with a sign that warns of the firm's pending doom as a prototypical example of this class of stakeholder.

When introducing the general category of stakeholders who possess not one, but two of the three stakeholder attributes described above, Mitchell *et al.* make an interesting proposition. In essence, they suggest that when entities possess any two stakeholder attributes, they progress from a passive to an active disposition and that this transformation embodies the vehicle of their increase in relevance with respect to stakeholder selection. Accordingly, the first such class of stakeholder, those that possess not only legitimate claims, but also the ability (*i.e.* the power) to act on those claims, called dominant stakeholders, is argued to be of significant importance to a firm as they constitute what is called its *dominant coalition* [123].

Mitchell *et al.* call the next class of stakeholders dependent stakeholders, those with urgent and legitimate claims, but not the power to pursue those claims. Interestingly, despite being characterised by the inability to carry out their will independently, the authors urge managers to consider the concerns of dependent stakeholders actively as they may, and often do, succeed in gaining the advocacy of more powerful stakeholders [325]. The second to last class of stakeholder, called dangerous stakeholders, is rather an unpopular sort of entity. Mitchell *et al.* [325] describe this class of stakeholder (those characterised by considerable power and urgency, but not legitimacy) as the kind of entity that may employ coercive and possibly dangerous means to influence a firm. The advice given in respect of how managers should go about dealing with

such an entity is two-sided. On the one hand, the authors criticise this class of stakeholder, stating that they are uncomfortable with the idea that dangerous stakeholders be afforded a measure of legitimacy on the back of the typology they propose, but at the same time, the authors state that a failure to identify dangerous stakeholders may result in the more disastrous failure of neglecting to counteract their influence [325]. Accordingly, Mitchell *et al.* suggest that dangerous stakeholders be considered, but with the sole purpose of ensuring the mitigation of their cause. Not much need be said regarding the final class of stakeholder, called definitive stakeholders as they are characterised by the entire spectrum of stakeholder attributes and are, as such, per definition, stakeholders.

Since its publication, the work of Mitchell *et al.* [325] has become one of the most cited contributions related to the field of stakeholder analysis[3]. Accordingly, it is employed as a foundational piece of literature with respect to the work presented later in this dissertation.

8.3 Making strategy: The journey of strategic management

Methodologies dedicated to the classification of stakeholders in the context of corporate endeavour can typically be grouped into one of two general categories: Those that focus an analyst's attention upon entities that have an interest in, or are affected by, the strategic future of a firm, regardless of whether or not those entities possess significant power, and those that regard only the powerful [141]. The next methodology to be discussed, certainly belongs to the latter of these categories. In fact, during its introduction, the authors, Eden and Ackerman, explicitly state that the method they propose is strictly intended to facilitate what they argue to be the *utilitarian* aim of identifying entities that have, or can develop, the power either to *support* or to *sabotage* the successful management of a firm[4]. *Power* is thus clearly employed as the governing determinant of stakeholder relevance. Interestingly, however, in the actual method proposed by Eden and Ackerman, the authors employ the notion of *interest* as a second dimension of stakeholder relevance in order to construct a two-dimensional categorisation of potential stakeholders. More specifically, Eden and Ackerman employ the following dimensions to characterise the importance of an entity as stakeholder to a firm: An entity's level of interest in the strategic activity of the firm, and the measure of its power to influence the achievement of the firm's strategic intent [4].

Based on the aforementioned dimensions, the authors proceeded to define four unique stakeholder classes, namely *players*, representing affected actors that possess both the intent and the power to influence the strategic future of a firm, *context setters*, representing entities characterised by the passive possession of considerable power, *subjects*, representing entities that would like to, but do not possess the power to, influence the strategic future of the firm, and finally, what is called the *crowd*, representing the unaffected bystander that is of little to no interest to the firm. These stakeholder classes, when put together, form what Eden and Ackerman call the *power/interest grid*, are depicted in Figure 8.2.

As described by Eden and Ackerman, the main purpose of the above categorisation is to aid analysts in the process of separating stakeholders that are relevant to a firm's strategic future from those that merely constitute the periphery. The process of identifying a collection of potential stakeholders from which such key stakeholders may be identified is, however, never addressed

[3]The article title *Toward a theory for stakeholder identification and salience: Defining the principle of who and what really counts* is credited with over 12 000 citations on Google Scholar [325].

[4]The utilitarian approach employed by Eden and Ackerman is certainly in harmony with the purpose that seems to underlie the method they propose: The identification of key stakeholders in order to manage the strategic alliances among them.

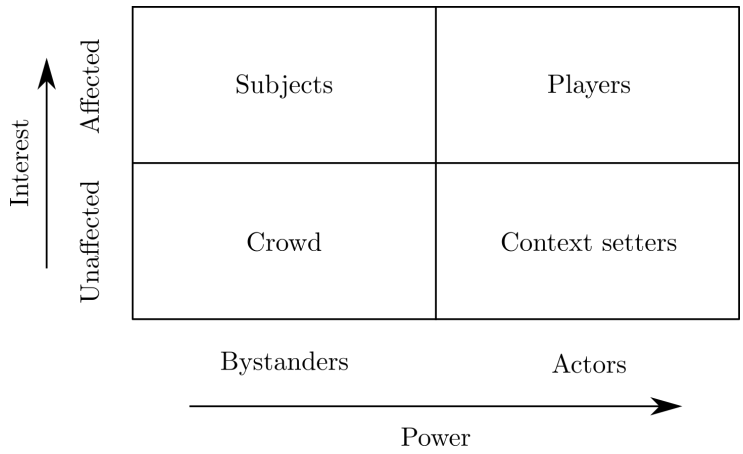


FIGURE 8.2: *A classification of stakeholders according to the dimensions of power and interest, called the power/interest grid [141].*

by Eden and Ackerman (which is why the method they propose belongs to the second category of the typology proposed by Reed *et al.* [376]) [141]. After introducing the power/interest grid, the authors proceed to discuss the next phase in the methodology they propose, called the *star diagram*. Before proceeding to a discussion of the star diagram, there is one final component to the power/interest grid that should be noted. Eden and Ackerman remind analysts that in addition to determining to which class a stakeholder belongs, it is necessary to take note of the unique disposition of each stakeholder [4]. More specifically, the authors suggest that in addition to determining the extent of a stakeholder's power and interest, analysts would do well to note whether the stakeholder is positively, negatively or variably disposed to the strategic future of a firm[5].

When introducing the star diagram, Eden and Ackerman describe how the application of the aforementioned power/interest grid may lead to somewhat biased results, because of unquestioned assumptions related to its application. Accordingly, the star diagram is presented as an attempt at mitigating such bias. To elucidate, the star diagram, depicted in Figure 8.3, is designed to facilitate the process by which analysts may attempt to determine (1) the basis of a stakeholder's power, and (2) the basis of its interest, thereby making salient and verifying any assumptions related thereto[6]. The construction of a star diagram consists of the following steps: First, an analyst is required to describe the manner in which stakeholders perceive and respond to a firm and its planned future endeavours (*i.e.* its spectacles). Thereafter, he or she must proceed to describe the manner in which the strategic future of the firm affects the goals pursued by said stakeholders (*i.e.* its affected aspirations). Thirdly, the sources of a stakeholder's power (*i.e.* its support mechanisms) should be made salient. Finally, the different ways in which a stakeholder may select to use its power should be described (*i.e.* its available sanctions[7]) [2].

[5]Eden and Ackerman [141] suggest that analysts indicate the determined disposition of a stakeholder in a colour-coded fashion within the power/interest grid using the following colour scheme: Potentially hostile stakeholders should be coloured red, potentially collaborative stakeholders should be coloured green, and finally, stakeholders whose disposition is different with respect to different parts of a firm's strategic future should be in blue.

[6]As the process of constructing a star diagram can be time-consuming, Eden and Ackerman [141] suggest only doing so for stakeholders determined to be most important to the strategic future of a firm during the prior application of the power/interest grid.

[7]Interestingly, Eden and Ackerman [141] do not describe what they mean by available sanctions in great detail and, as such, the above description thereof is based, in part, on the author's interpretation.

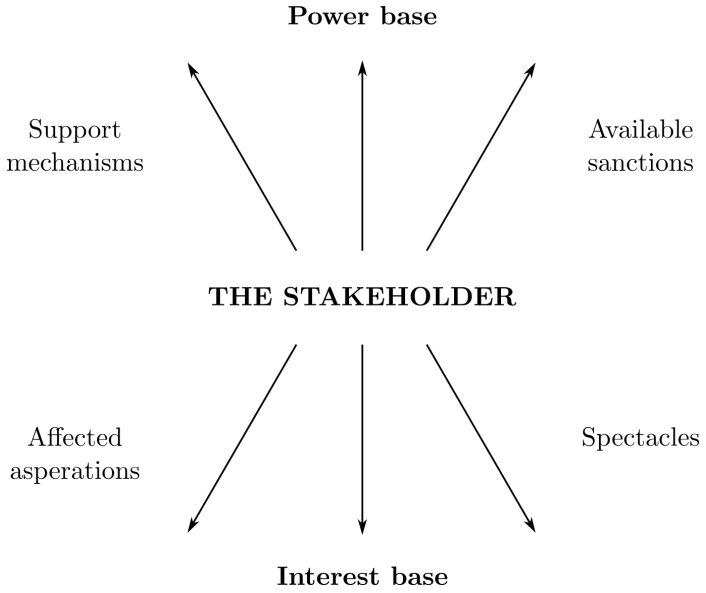


FIGURE 8.3: *A star diagram designed to facilitate the elicitation of the bases of both an entity's power and its interest* [141].

Eden and Ackerman proceed to discuss what they call *actor influence network maps*. Essentially, the authors argue that analysts should, at this point in an analysis, consider the relationships between potential stakeholders as doing so may elicit the central 'conduits' of power within a select group of stakeholders. The authors do not, however, present a method for facilitating such an analysis and, as such, the account of Eden and Ackerman's work is concluded [141].

8.4 A stakeholder circle methodology

Similar to the methods described in §8.2 and §8.3, the *stakeholder circle methodology*, proposed by Bourne in her doctoral dissertation titled *Project relationship management and the stakeholder circle* [69], is devised to support analysts in assessing the relative importance of a set of pre-identified stakeholders in order to facilitate their selection. The methodology essentially consists of a two-phase classification process, one phase dedicated to the classification of potential stakeholders according to the measure of their *power*, *proximity* and *urgency*, and the other one dedicated to the classification of potential stakeholder according to both the level of their *interest*, and the nature of their *disposition* [69]. The elements involved in the execution of these phases are described below.

To open the discussion on Phase 1, Bourne plainly defines the term *power* to represent the extent to which an entity possesses the ability to "kill a project," and then simply proceeds to state that analysts are to characterise this notion on a scale from 1–4, where 1 implies that an entity has relatively little power and 4 implies that an entity has a definite capacity to formally enforce change [69]. Thereafter, Bourne proceeds to present an equally concise characterisation of the measure of an entity's proximity[8], stating that, like power, analysts should represent this notion on a scale from 1–4, where 1 implies that an entity is relatively removed from a project, and 4 implies that an entity is directly involved therein. The final dimension of stakeholder relevance employed during the method's initial phase is that of *urgency*. To define this dimension formally, Bourne adopts a definition proposed by Mitchell *et al.* [325] in an article titled *Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts*

[8]The notion of proximity is never formally defined by Bourne as she considers it self-explanatory [69].

(see §8.2 for a discussion thereof). Interestingly, in the discussion on the manner in which analysts are to characterise an entity's urgency, they are instructed to do so not by the actual measure thereof, but by the affect of that measure upon the manner in which analysts are forced to respond to the related entity's concerns [69]. More specifically, analysts are instructed to rate the degree of an entity's urgency on a scale from 1–5, where 1 implies that analysts need not pay much attention to an entity's concerns, and 5 implies that immediate action is warranted, irrespective of other commitments [69].

The second phase of Bourne's method is described in no more detail than the first. It is presented plainly to consist of the execution of two separate analyses, one to elicit the level of an entity's interest in a project, and one to elicit the nature of the entity's disposition toward it [69]. Bourne neglects to provide any guidelines according to which analysts are to go about conducting these analyses, simply stating that the level of an entity's interest should be rated on a scale from 1–5, where 1 implies *antagonism*, 3 *ambivalence*, and 5 *commitment*[9], and similarly, that an entity's disposition should be rated on a scale of 1–5, where 1 implies *active support*, 3 represents a *non-committal* entity and 5 implies *active-opposition* [69].

After introducing the phases described above, Bourne proceeds to present what she calls a "complex set of Excel spreadsheets" capable of facilitating the visualisation of a potential stakeholder's relevance. To elucidate, Bourne developed a collection of Excel spreadsheets that, based upon the aforementioned ratings regarding the measure of an entity's power, proximity, urgency, interest and disposition, is capable of producing as output what is called a *stakeholder circle* [69]. For the sake of illustration, a hypothetical example of a stakeholder circle in respect of eleven unique stakeholder groups is depicted in Figure 8.4. The figure may be interpreted according to the following guidelines: The number of concentric circles between the centre and the shape representing a stakeholder symbolises the proximity of that stakeholder to the related project. The more uniform the colour-scheme of the shape representing a stakeholder, the more homogeneous the views of the individuals that constitute that stakeholder[10]. And finally, the size of the shape representing a stakeholder indicates the extent of its power. As can be seen, there are several dimensions of stakeholder relevance that formed part of the method proposed by Bourne that are not incorporated in the stakeholder circle (*e.g.* urgency, interest and disposition). Bourne, however, never addresses this lack of representation.

In conclusion, although this chapter would be incomplete without a review of the stakeholder circle methodology proposed by Bourne, it is unclear why the method has received attention in the academic literature. The method not only leaves much unaddressed, but fails to present the methodological components that are addressed in such a way that their working and sociological underpinnings are adequately elucidated. All in all, the methodology proposed by Bourne is hard pressed when compared to some of the methods discussed in the previous sections.

8.5 How to do, or not to do, a stakeholder analysis

Varvasovszky and Brugha open the account of their method for the classification of stakeholders titled *A stakeholder analysis* [465] with the following description of the general purpose thereof: A stakeholder analysis represents an attempt at generating knowledge about a select group of actors toward understanding their behaviour, interrelations and/or interests in order to assess the general measure of the influence that they bring to bear on the decisions to be made during

[9] The scale suggested to characterise an entity's interest is above all else confusing. Not only do the labels suggested (*i.e.* antagonistic, ambivalent and committed) relate poorly to the notion of interest, the scales proposed to characterise the notions of interest and disposition are essentially one and the same.

[10] Bourne neglects to provide a formal definition for the notion of a stakeholder [69].

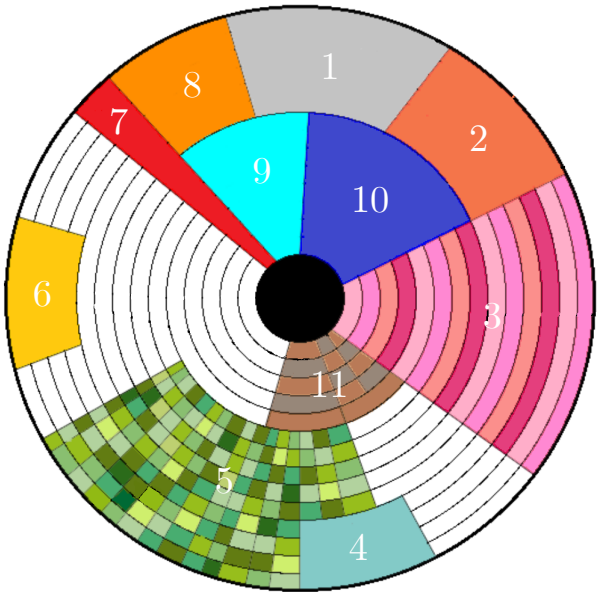


FIGURE 8.4: *A hypothetical example of the output of the methodology proposed by Bourne [69], called a stakeholder circle. The figure may be interpreted as follows: The number of concentric circles between the centre and the shape representing a stakeholder symbolises the proximity of that stakeholder to the related project. The more uniform the colour-scheme of the shape representing a stakeholder, the more homogeneous the views of the individuals that constitute that stakeholder. And finally, the size of the shape representing a stakeholder indicates the extent of their power [69].*

an endeavour [465]. This rather elegant description of the purpose of conducting a stakeholder analysis is in many ways representative of the manner in which Varvasovszky and Brugha go about presenting the method they propose — detailed to the point of being comprehensive.

The method is introduced with what the authors call a set of *preliminary questions* seemingly designed to prime the pending analysis. To elucidate, Varvasovszky and Brugha present three such questions, phrased as follows: "What is the aim and time dimension of the analysis?" "What is the context of the analysis?" And finally, "at what level will the analysis take place?" The dynamics these questions are designed to address, as made salient in the discussion related to their introduction, are depicted in Table 8.1 [465]. As can be seen, a total of six dynamics are listed in the left-hand column of the table, while the respective options associated with the consideration of each dynamic are listed in the right-hand column. Analysts are instructed to determine (1) whether the *purpose* of their analysis is aimed at *evaluation*, *implementation*, or *planning*, (2) whether the *focus* of their analysis is *retrospective* or *prospective* in nature, (3) whether they are *interested* in the *process* of design, or simply the *objectives* being pursued, (4) whether the *scope* of their analysis is *broad* or *narrow* in focus, (5) whether the *time-frame* of the analysis is *rapid*, *short-term*, *medium-term* or *long-term*, and (6) whether the *stage* at which the analysis takes place is *pre-implementation*, during *planning*, *development* or *post-implementation*[11] [465]. Following these preliminary considerations, Varvasovszky and Brugha proceed to discuss the actual analysis of potential stakeholders.

The discussion on the analysis of stakeholders opens with an account of the dimensions that the authors argue underlie the notion of stakeholder relevance, namely the extent of their *influence*, the level of their *interest*, and the nature of their *disposition*. Unfortunately, the authors neglect

[11] The authors include the notion of an *historical analysis* in respect of the stage at which an analysis may take place. This is, however, neglected as its inclusion is illogical — the stage of an analysis cannot be historical as this would imply that the analysis has already taken place.

TABLE 8.1: *A list of the dynamics analysts are to consider during the preliminary phases of a stakeholder analysis, adapted from Varvasovszky and Brugha [465].*

Dynamic	Options
Purpose	Evaluation
	Implementation
	Planning
Focus	Retrospective
	Prospective
Interest	Process
	Objective
Scope	Broad
	Narrow
Time frame	Rapid
	Short-term
	Medium-term
	Long term
Stage	Historical analysis
	Pre-implementation
	Planning
	Development
	Post-implementation

to define these dimensions formally, although they do proceed to discuss a related dynamic thereto not noted in any other method reviewed in this chapter — the consideration of how the qualities of an entity may change over time [465]. Consider Figure 8.5, which relates to a case study conducted by Varvasovszky and Brugha in the alcoholic beverage sector of a country not named and depicts the manner in which the influence and disposition of four key stakeholders may change over time. The following stakeholders are considered: The *Association of Spirit Producers* (ASP), the *National Public Health Institute* (NPHI), the *Transport and Road Safety Division of the Police* (TRSDP) and the *National Institute for Alcohol* (NIA). Although the value of such an analysis is multi-faceted, the authors suggest that it is most relevant to the identification of key future role players with whom strategic partnerships may be pursued [465]. Following the discussion of this time-based analysis of stakeholders, Varvasovszky and Brugha proceed to introduce the final component of their method — one dedicated to the selection of an optimal strategy for the management of stakeholders.

Although not explicitly stated, it seems that the final phase of the method builds collectively upon the work done in those before it. The working of this component of the proposed method is best illustrated visually and, to that end, consider Figure 8.6. As can be seen in the figure, the optimal strategy for managing stakeholders is determined according to the *position* that characterises them (*i.e. supportive*, *mixed*, *non-supportive* and *marginal*) which, in turn, is determined by the measure of a stakeholder's influence and the nature of its disposition [465]. More specifically, the proposed "optimal fit" between the position of a stakeholder and the strategy proposed for its management is to *involve* supportive stakeholders, *defend* against non-supportive ones, *collaborate* with "mixed-blessing" stakeholders (*i.e.* those who combine elements of support and non-support), and finally, to *monitor* marginal stakeholders (*i.e.* those that exhibit a low level of threat and limited potential for collaboration) [465].

In a manner that alludes to the maturity with which Varvasovszky and Brugha approach the presentation of their work, the account of their method is concluded with a discussion on the

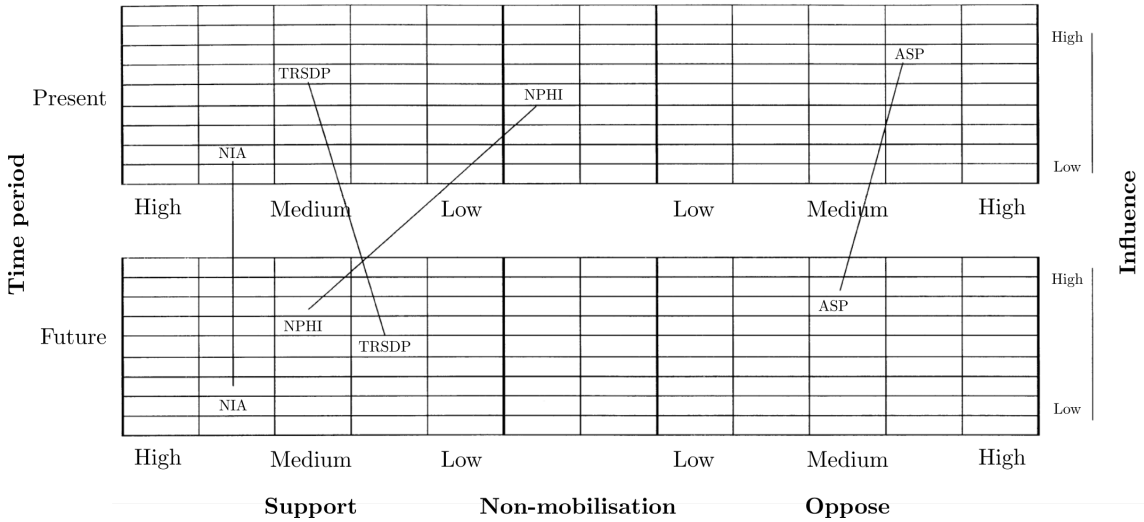


FIGURE 8.5: *A hypothetical output of an analysis of how stakeholders' influence and disposition may change over time, conducted by Varvasovszky and Brugha in a case study in the alcoholic beverage sector of an unspecific country [465].*

Strategies Positions	**Involve**	**Collaborate**	**Defend**	**Monitor**
Supportive	Optimal fit	Missed opportunities	Missed opportunities	Missed opportunities
Mixed	Risk	Optimal fit	Missed opportunities	Missed opportunities and Risk
Non-supportive	Risk	Risk	Optimal fit	Risk
Marginal	Resource waste	Resource waste	Resource waste	Optimal fit

Optimal fit between diagnosed position and strategy

Suboptimal fit leading to missed opportunities for gaining support

Suboptimal fit leading to excess attention to low potential stakeholders

Suboptimal fit placing venture/ organisation at risk

FIGURE 8.6: *A grid illustrating the alternative strategies for managing dissimilar groups of stakeholders, and the consequence thereof [465].*

limitations thereof. To elucidate, the authors make salient that any stakeholder analysis, and so too their method, is fundamentally *cross-sectional* in nature as it provides but a snapshot of what may be a rapidly changing context [76]. This quality is portrayed as a potential weakness as it implies that stakeholder analyses are fundamentally influenced and coloured in by the real-time elements of their context, including the stakeholder analysis itself. In essence, Varvasovszky and Brugha warn that analysts should remain weary of the potential bias in any stakeholder analysis and not fall victim to the trap of finding unwarranted comfort in the guidance of a method that is simply designed to fulfil the role of support [465].

8.6 Stakeholder analysis combined with social network analysis

In the multi-dimensional kind of endeavour that often relates to the context of environment policy making, stakeholder analysis has a well-worn seat at the table. It is thus rather appropriate that the final methodology to be discussed, belonging to the second and third categories of the typology proposed by Reed *et al.* [376], is one specifically designed for such a more involved kind of context. Proposed by Lienert *et al.* [288] in an article titled *Stakeholder analysis combined with social network analysis provides fine-grained insights into water infrastructure planning processes*, the methodology addresses both the classification of potential stakeholders and the analysis of the relationships between them. More specifically, the methodology is a two-phase process, the first of which, called *stakeholder analysis*, is dedicated to the categorisation of stakeholders, and the second, called *social network analysis*, is dedicated to an analysis of the relationships among stakeholders [288]. In the remainder of this section, these phases are discussed in turn, starting with the latter.

There is not much that need be said in respect of the categorisation method presented by Lienert *et al.* as it is remarkably similar to that contained in the methodology proposed by Eden and Ackerman [141] (discussed in §8.3). They too, employ the notions of power (*i.e.* the ability to influence) and of effect to characterise dissimilar stakeholders along a two-dimensional spectrum of relevance. The only real difference is that Lienert *et al.* select to discretise the spectrum of stakeholder relevance according to a more refined classification (adopted from Likert [289]). More specifically, the measure of an entity's influence is rated as being either 0, 2.5, 5, 7.5 or 10, representing an entity that has no influence, has very limited influence, has influence, has strong influence, and finally, has governing influence, respectively. Similarly, the degree to which an entity is affected by an endeavour is rated as being either 0, 2.5, 5, 7.5 or 10, representing an entity that is slightly affected, is affected, is strongly affected, and finally, is very strongly affected, respectively [288]. The remainder of phase 1 of the methodology proposed by Lienert *et al.* is simply a carbon copy of the classification proposed by Eden and Ackerman [4] and, as such, the intricacies thereof are not re-documented here.

Social network analysis, as defined by Lienert *et al.* [288], essentially represents a method designed to support analysts in uncovering the system of relationships among the actors of a network such that the *central* players therein are made salient. The authors describe the groundwork related to such an analysis to consist of first writing the names of potential stakeholders (as identified during phase 1) onto small cards and then arranging those cards on a large piece of paper such that the system of relationships among stakeholders can be identified and drawn in using coloured markers[12] [288]. Once the network of relationships has been constructed[14], Lienert *et al.* state that analysts may determine both the *structural importance* and the *power*[15] of stakeholders by calculating each stakeholder's *degree centrality* and *betweenness centrality*, respectively. Unfortunately, not much detail is given in respect of the criteria that underlie these notions, the authors simply state that degree centrality relates to the ties that an actors shares directly with other actors (whether the number or nature of such ties are of importance is never clarified), while betweenness centrality is described to represent the number of times an entity

[12] The authors list four unique kinds of relationships with respect to which such an analysis should be conducted (represented in dissimilar colours), namely *cooperative* relationships, *financial exchange* relationships, *pressure* relationships, and *conflict* relationships[13] [288]. By conducting these analyses, Lienert *et al.* argue that analysts will be able to elicit the central players within the network of stakeholders related to a problem context.

[14] Lienert *et al.* [288] do not provide further guidance as to how analysts are supposed to uncover such relationships.

[15] Since the notion of power was employed to classify potential stakeholders during phase one of the methodology they propose, the fact that Lienert *et al.* employ the term here is rather confusing, especially since they neglect to define what is meant thereby.

is present along shortest indirect paths between two, non-adjacent entities [288]. By conducting these analyses, Lienert *et al.* argue that analysts will be able to elicit the central players within a network of stakeholders related to a problem under consideration. To conclude the discussion, consider a prototypical example of a social network analysis, adapted from Lienert *et al.* [288] and depicted in Figure 8.7.

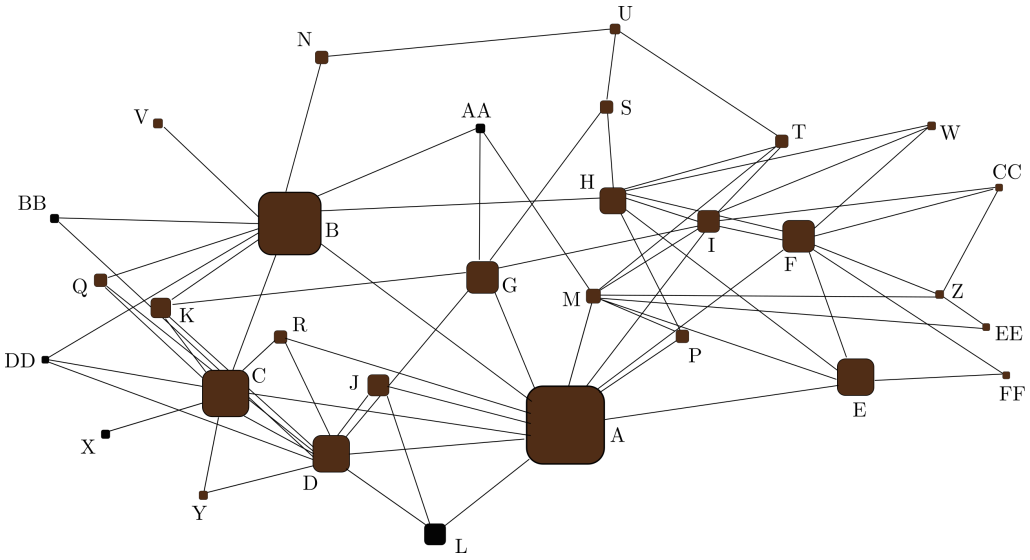


FIGURE 8.7: *A stereotypical example of the output of a social network analysis, where the nodes labelled A–GG represent stakeholders, and the size of each node indicates the degree of betweenness centrality related to the stakeholder it represents, adapted from Lienert* et al. *[288].*

8.7 A high-level comparative analysis of the methods discussed

The methods discussed in the preceding sections are held to be representative of the academic literature related to stakeholder theory. Toward presenting the reader with as detailed a review of these methods as appropriate, Table 8.3 is dedicated to an illustration of their distinction according to the features that characterise them. In general, the methods considered[16] are distinguished according to (1) their designed purpose, and (2) the dimensions of stakeholder relevance they employ. To facilitate its presentation, Table 8.3 is constructed according to the legend depicted in Table 8.2.

Before concluding this chapter with a brief summary, consider the following shortcoming related to the methods reviewed therein that utilise the notion of power to distinguish relevant from irrelevant entities. All but one fail to provide a working description of the components that constitute power. Given that there is a plethora of definitions for the notion of power in the literature, this neglect is rather odd. To illustrate the potential utility of such definitions to the task of stakeholder selection, the work of psychologists French and Raven [174], and of economist Galbraith [179] is briefly reviewed below[17].

French and Raven [174] defined the notion of power according to what they called the five bases of social power, namely *legitimate power*, *reward power*, *coercive power*, *expert power*,

[16] It should be noted that only feature-based methods, as opposed to role-based methods, are considered. Role-based methods are considered inadequate in the context of the work presented in this dissertation [366].

[17] For a more general review of contributions made toward defining the notion of power, see De-Moll's *Everyday experiences of power* [124].

TABLE 8.2: *The legend employed to represent the dimensions of stakeholder relevance related to dissimilar stakeholder selection methods with a view to facilitate a comparative analysis of the methods in Table 8.3.*

Label	Dimension
1	Power
2	Legitimacy
3	Urgency
4	Affect/Interest
5	Disposition
6	Convenience
7	Proximity
8	Subjective judgement
9	Degree centrality
10	Betweenness centrality
11	Not addressed

TABLE 8.3: *A comparative analysis of a collection of stakeholder selection methods according to the features that distinguish them. In cases where methods do not have uniquely identifiable names, the related authors' names are listed.*

Method	Purpose			Dimensions governing stakeholder relevance										
	Identify	Classify	Analyse	1	2	3	4	5	6	7	8	9	10	11
Focus group [173]	×													×
Semi-structured interview [136, 296]	×													×
Snowball sampling [64, 75]	×													×
Purposive sampling [147]	×										×			
Convenience sampling [155]	×								×					
Mitchell *et al.* [325]		×		×	×	×								
Eden and Ackerman [4]		×		×			×	×						
Lienert *et al.* [288]		×	×	×			×	×				×	×	
Varvasovszky and Brugha [465]		×		×			×	×						
Bourne [69]		×		×				×		×				

and *referent power*. These so-called power bases were defined as follows: Legitimate power represents any form of influence that stems from the internalised values of individuals dictating that one individual has certain rights over another. Reward power embodies influence granted an individual because of its ability to induce positive, and/or terminating negative experiences. Coercive power, on the other hand, addresses the influence granted an individual as a result of its ability to induce negative, or terminating positive experiences. Expert power is defined to represent the influence granted an individual because of its knowledge and/or skill in a certain area of interest, and finally, referent power is granted an individual because of its standing with other, third party individuals [174].

Galbraith [179] proposed an alternative definition for the notion of power that addresses not only the *sources* from which power originates (as French and Raven [174] had done), but also the *forms* in which such power may be exerted. Galbraith proposed three sources of power, namely the *personality* of an individual, the *property* it possesses, and the *organisational* position it holds, and three forms of exertion, namely *condign*, *compensatory*, or *conditioned* power [179]. In essence, the first of the sources of power above is defined to represent an individual's intelligence, dominance and ability to inspire, the second simply represents the material resources possessed by an entity, the third source of power seems to have been considered self-explanatory by Galbraith as he neglected to formally describe what it represents [179]. In a similar fashion, the forms of power proposed were only briefly described. In the order listed, they were defined to represent power that is *based on force*, power that results from an *exchange of resources*, and power that is a *product of persuasion* [179].

Evidently, the work of researchers such as French and Raven [174] and of Galbraith [179] represents an opportune foundation from which to develop a definition for the notion of power tailored to the task of stakeholder selection. It is unfortunate that contributions such as these were not drawn upon during the development of the methods discussed in this chapter.

8.8 Chapter summary

This chapter was dedicated to a review of the academic literature related to methods designed for stakeholder analysis. In order to facilitate their discussion, these methods were reviewed in terms of a typology proposed by Reed *et al.* [376] for classifying methods into one of three categories according to whether they were designed to facilitate (1) the identification of stakeholders, (2) the classification of stakeholders, or (3) the analysis of the network of relationships among stakeholders. More specifically, §8.1 was dedicated to a review of methods belonging to the first of these categories, §8.2–§8.5 to methods belonging to the second, and §8.6 to methods belonging to the third of the aforementioned categories. In addition, §8.7 contained a brief analysis of several stakeholder methods according to the features that distinguish them.

Based on the work reviewed in this chapter, it is clear that there is much to be desired in the field of stakeholder selection theory. The lack of methodological maturity is the most obvious in respect of methods belonging to the first of the aforementioned categories (*i.e.* those designed to facilitate the identification of stakeholders). Accordingly, the work presented later in this dissertation represents, in part, an attempt at taking the field of stakeholder analysis at least one step closer to maturity.

Part II

The meta-methodology for real-world problem solving

As delineated in Chapter 2, the *meta*-methodology proposed in this dissertation is developed in accordance with the design framework proposed by Herman [211]. As such, Part II is structured in the following way: In Chapter 9, a preliminary version of the proposed methodology is presented. Thereafter, the application of the methodology to a theoretical case study is presented in Chapter 10, along with several modifications thereto deemed necessary on the back of its application. In Chapter 11, feedback received from an expert in the application of qualitative methodologies of the kind proposed, is reviewed and incorporated into a final version of the methodology. This final version of the methodology is then applied in Chapter 12 to a practical case study aimed at illustrating its utility. To conclude Part II, an abstract mathematical model developed in an attempt to uncover dynamics related to the context for which the proposed methodology was designed, but which are typically hidden from a purely qualitative analysis, is presented in Chapter 13.

CHAPTER 9

A systems solution meta-methodology

Contents

This chapter is dedicated to the explication of a meta-methodology designed to facilitate the solution of real-world problems. A high-level overview of the proposed framework is presented in §9.1, and this is followed by an in-depth discussion on the various phases of the meta-methodology in §9.2–§9.3. Thereafter, the chapter is brought to a close with a short summary in §9.4.

9.1 A general depiction of the methodology

In order to stand upon the shoulders of academic giants gone before, it is necessary first to delve into their accomplishments. The literature review conducted toward that end in Part I is extensive, both in nature and purpose. Despite this extensive quality, it is possible to capture concisely the essence of what was discussed as the intricacies of human judgement, the intricacies of real-world problems, and the technicalities involved in the working of solution methodologies for accommodating the interaction of these phenomena during the process of real-world problem solving. In harmony with the philosophy of the critical systems movement, discussed in §6.8, it has been argued that in order to develop a mature solution methodology, neither the human, nor the technical elements of the real-world can be neglected — that it is only by synthesising these phenomena that one can develop a response to real-world problems which is able to accommodate its intricately nuanced nature. Indeed, the framework for real-world problem solving proposed in this dissertation has more in common with the critical systems movement than any other paradigm, although the critical systems movement does not contain it as its commitment toward human emancipation is somewhat neglected.

Before embarking on an exposition of the proposed methodology, it is necessary to describe the context of application for which it was designed[1]. Simply put, the proposed methodology is designed to address the kind of real-world problems that exist in systems exhibiting the following property: The entities contained therein are forced to collaborate despite pursuing conflicting

[1]The proposed methodology may well be of utility to address a wider scope of problems than those for which it has been designed.

outcomes (*i.e.* the outcomes pursued by competing entities are conflicting, but interdependent). As a result of the intertwined nature of this kind of system[2], the problems contained therein are often exceptionally difficult to solve as it is never clear who to consult in order to confirm the appropriateness of a proposed solution (*i.e.* it is unclear who the stakeholders of the problem are) and, furthermore, the process of problem solving[3] entails not only the identification of a solution that is technically appropriate, but one that simultaneously distributes the benefit to be attained by relevant stakeholders appropriately. Because of these qualities, the activities that dominate the methodological process of solving these kinds of problems include: The identification of relevant stakeholders, the determination of outcomes pursued, the prioritisation of selected outcomes, the establishment of aspiration values, and the selection of an appropriate solution. It is toward facilitating these activities, whereby real-world problems of the kind described above can be solved, that the meta-methodology proposed in this chapter is aimed.

In general, the proposed methodology is partitioned into three phases[4], namely *problem formulation*, *response formulation* and *response implementation* as depicted in Figure 9.1. These phases represent the process of developing a formal representation of a problematic situation that adequately facilitates the pursuit of an appropriate response to the problem considered, the process of actually developing an appropriate response, and the process of practically implementing that response, respectively. As illustrated in Chapters 6 and 8, many academics have contributed to the collective pursuit of developing methodologies for managing these components of real-world problem solving (*e.g.* the methods of the systems movement). The framework proposed in this dissertation is not an attempt at reinventing the wheel, but is, where possible, built upon the work covered in the literature review. Before preceding, it should be noted that although the tasks of response formulation and response implementation are included in the high-level overview of the proposed methodology below, their specific design is not addressed in this dissertation as it is considered outside the scope of this work, as noted in §1.4.

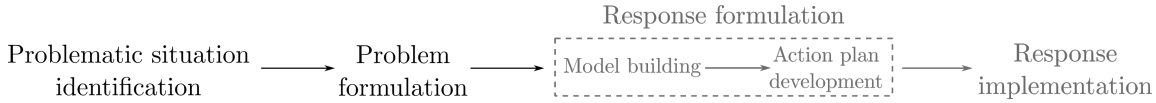


FIGURE 9.1: *A high-level overview of the phases within the proposed methodology.*

9.2 Problem formulation

The manner in which Figure 9.1 has been designed may result in the unintended impression that the identification of a problematic situation constitutes the first phase within an *endeavour*[5]. In reality, however, what is being implied is simply the acknowledgement of some non-ideal situation, and that something needs to be done to address it. In other words, the initial identification of a problematic situation is a naturally occurring event that takes place outside the context of an endeavour and establishes the need for an endeavour towards addressing this non-ideal situation. The first phase of this methodology is thus that of *problem formulation* — a phase characterised by considerable complexity. In the sections that follow, the components proposed to constitute

[2]One can describe this kind of system by stating that such entities are linked by a complex network of interdependent pursuits, activities and resources in which each entity possesses varying degrees of power and autonomy.

[3]The problems that surface in this kind of system typically belong to the pluralist and coercive categories of the problem typology discussed in §7.4.

[4]The practical distinction between these phases is at times rather ambiguous.

[5]A term representative of the system in which the process of real-world problem solving is contained.

the process of problem formulation, namely *stakeholder selection* and *problem structuring*, as well as a feature that relates to all the phases of the methodology, called the *decision-making checklist*, are depicted in Figure 9.2 and are discussed in detail in this chapter.

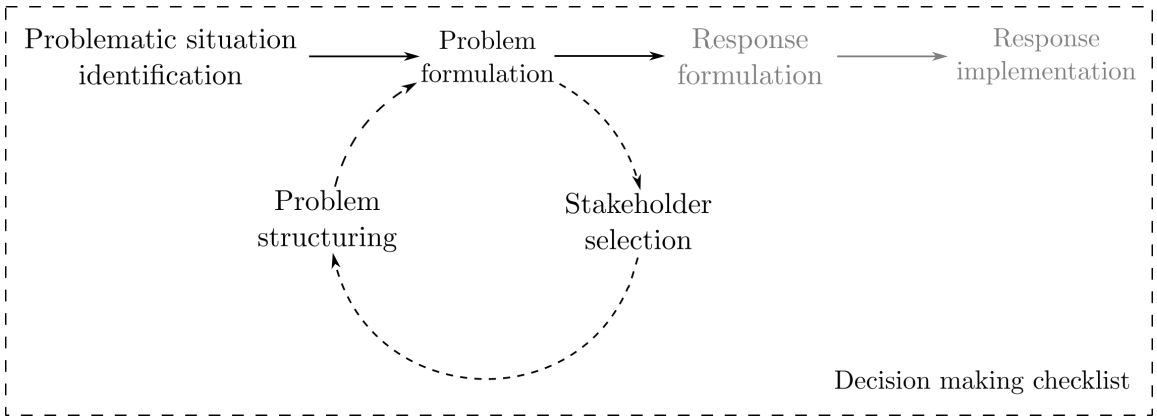


FIGURE 9.2: *A high-level overview of the phases within the proposed methodology together with a lower-level depiction of the phases that constitute the process of problem formulation.*

9.2.1 The identification of relevant stakeholders

Solution methodologies dedicated to the process of stakeholder selection are typically classified into one of three categories according to whether they are designed to aid analysts in identifying stakeholders, differentiating between and categorising stakeholders, or investigating the network of relationships among stakeholders. As made salient in §8.1 of Chapter 8, there is a gap in the academic literature concerning methods dedicated to the first of these priorities. In the words of Reed *et al.* [376], "Much of stakeholder analysis literature has presumed that stakeholders are self-evident and self-construed, and has focused on categorising pre-identified stakeholders to understand their interests and relationships. However, before this can be done, it is necessary to identify who holds a stake in the phenomenon under investigation." In an attempt to address this shortcoming, a methodology designed to facilitate both the identification and selection of stakeholders is proposed. To clarify, this methodology may be utilised as primer to methods dedicated to the second and third priority listed above (reviewed in Chapter 8), but it is designed to facilitate both the identification and selection of stakeholders independently.

As point of departure, Freeman's oft cited definition of a *stakeholder* is employed [172]. Freeman defined[6] stakeholders to be entities who may affect, or may be affected, by the outcome of an endeavour [172]. By recognising the building blocks that constitute this definition, two elements characteristic of potential stakeholders result. This is, the ability to influence the outcome of an endeavour (defined as a measure of *power* possessed), and the degree to which an entity is affected by an endeavour and the activities that surround it (defined as a measure of *effect* experienced). Note that these elements are subjective in nature as their description depends upon human perception. Consequently, it is necessary to provide a formal definition of the components that constitute the notions of power possessed and effect experienced.

Before attempting to define these elements, a distinction is made between two types of systems. In any endeavour there exists the real-world system that contains the identified problematic situation (referred to here as the *original system*), and then there is the system that contains

[6]This definition has been paraphrased to suit the context of this dissertation.

only certain elements of the original system deemed relevant to the problematic situation being addressed during an endeavour (referred to here as the *temporary system*). Within this classification, the original system contains every component of the real world that is involved in a problematic situation, while the temporary system contains only those components that are deemed relevant by the entities conducting an endeavour. With this in mind, the notions of power and effect are defined in the sections that follow.

The notion of power

As made salient in the literature review of Chapter 8, methodologies designed to facilitate the selection of stakeholders typically utilise some conception of power (among other features) to distinguish the relevant from the irrelevant. Few such methodologies, however, provide an adequate description of the dimensions that underlie the possession of power. Toward providing such a description tailored to the context of an endeavour, the work of the economist Galbraith [179] (among others) is drawn upon as inspiration[7].

To speak of power is to speak of the potential to influence the behaviour of another (*i.e.* the decisions made by another). In more precise terms, the phenomena that constitute the existence of such influence is defined along four dimensions, namely *form*, *source*, *potential* and *efficacy*. The first of these dimensions addresses the nature of exerted power and contains two sub-dimensions, namely *coercive* and *suasive* power. As the labels assigned to them suggest, coercive power represents the exertion of power by means of forcefully enacting one's will upon another, and suasive power represents the exertion of power by means of persuasion. The attentive reader will notice that coercive power is similar to what Galbraith [179] called condign power, and suasive power to the combination of what Galbraith [179] called compensatory and conditioned power.

The author does not believe it possible to typologise all sources of power contained in the mechanisms of influence. Consequently, no such claim is made. Instead, the strategy adopted is to attempt to typologise only the sources of power that are expected to dominate the power dynamics within the context of an endeavour. Sources of power are defined as either *formal* or *informal*, and then according to the sub-dimensions contained within each of these categories. Notably, formal elements of power represent the elements of power that are officially recognised and described within the formal structure of the original system, while informal elements of power represent the elements of power that are not officially recognised or described as sources of power, but do, in fact, contain the capacity to influence the outcome of an endeavour.

[7] There are two limitations to the description of power proposed by Galbraith [179] that render a novel definition necessary, one in respect of its structure, and the other in respect of its scope. Concerning its structure, there is an incoherence between the power source labelled *personality* and the form of employment labelled *compensatory*. In essence, Galbraith defined an individual's personality to include his or her level of intelligence, ability to inspire, and overall effectiveness, while the compensatory exertion of power was defined to represent any form of influence characterised by an exchange of resources. These descriptions allude to a discrepancy in Galbraith's formulation of power — it is difficult to imagine how an individual could utilise his or her personality in a compensatory form.

The second aspect of the work of Galbraith [179] that is contested pertains to the sources defined to contain the notion of power (*i.e.* personality, property, and organisational position). Not only is there a measure of interdependence among these categories, but they are also limited in their scope. Concerning the first issue, Galbraith elected to define an entity's personality as a separate source of power, thereby failing to realise that personality plays a decisive role in the exertion of any form of sourceable power. The suggestion that personality, or any element similar in meaning, be defined a unique source of power, is considered illogical. Concerning the issue of scope, the listed power sources are considered lacking. They fail, for example, to contain the manner in which the perceived, pre-empted actions of another may act as a source of influence.

In the context of an endeavour, the dominant elements of formal power are contained within two formal declarations of power. The first, more dominant source of formal power resides in the power structure that exists within an original system, called *structural power*. More specifically, this form of power originates from the formal description of the power relationships among entities according to their positions in the structure of the original system (not unlike Galbraith's organisational power dimension). The second source of formal power, inspired by the work of Suchman [438], is called *customary power*. Customary power represents the influence granted entities as a result of the official regulations of the society in which an endeavour is being conducted. An appropriate example of customary power can be found in the context of *sustainable development*. During the past few decades, the fact that the environment suffers irrevocable consequences because of the nature of human endeavours has become widely acknowledged [222]. In response to such intellectual developments, many countries have implemented regulating policies to ensure that human affairs are conducted within the domain of sustainability [39, 78, 199]. As such, it is often prescribed that an entity, representative of the sustainable mandate, is involved in decisions made in the context of an endeavour. On the other end of the spectrum, it is argued that the dominant sources of informal power in an endeavour are contained within the following elements: *status*, *necessity*, and *incapacitation*.

Of the above sources of informal power, the foremost element is rather similar to elements in Galbraith's work, while the two latter elements are quite unique. To elucidate, an entity's status represents the material and non-material commodities possessed by that entity, and is thus somewhat similar to the *property* dimension defined by Galbraith. An entity's necessity as a source of power, on the other hand, is entirely novel and represents its perceived necessity with respect to the purpose of the original system. In other words, it represents an entity's ability to inhibit the purpose of the original system of which it is a part by becoming passive in its participation. This dimension captures a wide variety of strategic moves within the power battle of an endeavour. Consider the case of *industrial action* as an example. It is argued that the source of power employed by the mechanisms of industrial action rests predominantly upon the necessity of those participating therein with respect to the purpose of the system at which the protest is aimed. The final source of power, appropriately labelled incapacitation, shares more than its novel nature with the notion of necessity. In fact, one could say that incapacitation and necessity are two sides of the same coin. Whereas necessity represents an entity's ability to inhibit the purpose of a system by becoming inactive, incapacitation represents an entity's ability to inhibit the purpose of a system by actively disrupting the mechanisms thereof (*e.g.* sabotage).

The third dimension of power, its potential, is defined simply as the magnitude of the measure of an entity's power with respect to each of the sources of power listed above — the greater its measure, the greater its potential. This dimension is established formally to ensure that practitioners remain sensitive to the source-specific measure of power entities possess as some power sources may be more consequential than others. A slightly more complex dimension, the efficacy[8] of power represents the mechanism by which the impact of a certain measure of employed power is determined. That is, an entity's level of skill in harnessing power. The influence of an entity's level of skill can be likened to the effect of a *foreign exchange* trader's expertise in trading one currency for another. In the world of forex trading, the ability to create value simply by knowing how and when to convert one currency into another is known as the ability to *play the market*. In like manner, although dependent upon a different skill set, an entity's skill in exerting power represents its capacity to "play the market" within the context of

[8]Note that the efficacy of power incorporates what Galbraith defined under *personality* as a source of power.

an endeavour. Stated more formally, the manner in which an entity goes about exerting its power may result in the effect thereof upon the outcomes of an endeavour being either disproportionally larger, smaller, or exactly proportional to the potential of its sourceable power.

Power has now been defined according to the four constituent dimensions, namely its form, source, potential and efficacy (see Figure 9.3 for a summary). It is argued that the collective scope of these dimensions sufficiently contain the mechanisms that dominate the playing field during an endeavour. That being said, there is, of course, more to the defined dimensions than discussed above. For the purpose of stakeholder selection, however, the current exposition is deemed appropriate. Later in this dissertation, the mechanisms of power will once again become the focus of the discussion, at which point the intricacies thereof will be explicated in more detail. As a precursor to this later discussion, there is one more concept that should perhaps be mentioned at this juncture — that of a *power-play*. Simply put, a power-play represents the practical exploitation of a specific source of power, in a specific form. In other words, seeing that five sources of power have been identified, together with their two potential forms of use, it follows that there are ten available power-plays. By employing this concept, it is possible to describe the potential and efficacy of power with respect to a specific power-play.

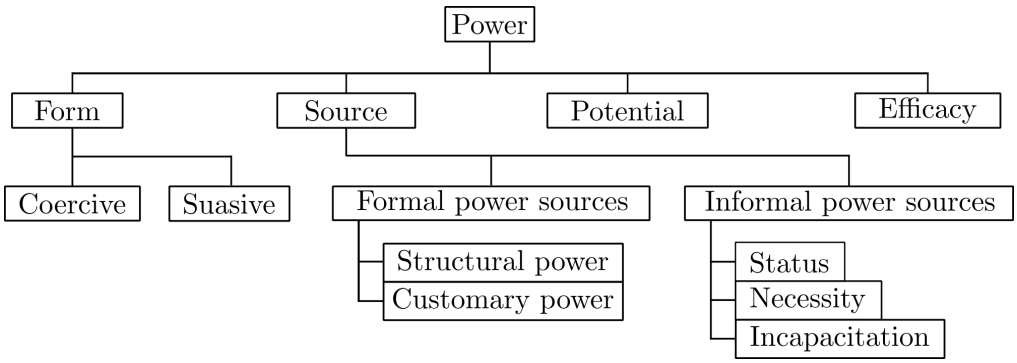


FIGURE 9.3: *An overview of the elements that dominate the mechanisms of power in an endeavour.*

The notion of effect

Establishing an exhaustive classification of the mechanisms by which entities may be affected by an endeavour has proven to be a persistent challenge in the academic literature [366]. It was Vos and Achterkamp [470] who argued that the identification of the affected within an endeavour poses a fundamental problem, as it is impossible to know when all such entities have been identified. In other words, no exhaustive set of categorical dimensions for the mechanisms of effect exist within the context of an endeavour. In a manner similar to the explication of power, no exhaustive description of elements contained within the mechanisms of effect is therefore given. Instead, the dynamics that dominate the mechanisms of effect are simply described in as much detail as appropriate[9]. More specifically, two dynamics are addressed: The criteria according to which the reality of effect experienced may be identified, and the mechanisms by which such effect propagates.

The notion of effect experienced is certainly a multifaceted concept. Yet, it is argued that the reality thereof may be uncovered according to the formal and informal outcomes that entities pursue in the original system in which an endeavour takes place, where the distinction

[9]Unlike the explication of power, no attempt is made here to define the dominant sources of effect — the mechanisms of human effect are so intricately entwined that it is not possible untangle them in a manner that adds value and maintains generality.

between formal and informal is made according to whether or not such outcomes are formally documented[10]. In other words, entities are considered affected if either the formal or informal outcomes they pursue relate in any way to the issues addressed by an endeavour, or the manner in which these issues are to be addressed. Once an entity is deemed affected, the severity of effect experienced may be determined according to the extent to which the outcomes it pursues are either simplified, or complicated by the issue under consideration. It is acknowledged that certain outcomes pursued, often formal in nature, deem an entity either affected or unaffected, without any shade of grade in between.

Describing the mechanisms by which the notion of effect exists is significantly more involved than simply stating the criteria according to which the severity thereof may be determined. To elucidate, it is necessary to explicate the very agency by which the effect caused by changes to the working of a system propagates. To that end, the concept of a *functional* relationship is introduced. To define this concept, a notion popularised by practitioners of an approach often adopted in the study of inter-organisational systems [122], called the *network approach*, is employed [366]. More specifically, the original systems containing an endeavour is considered to consist of a conglomeration of networks of interdependent *actors*, *resources*, and *activities*, where actors[11] are defined to represent human entities (*i.e.* in the form of individuals or collectives) who carry out goal-directed activities, which represent the tasks by which these actors are able to transform and transfer resources toward developing a more aggregated actor, and resources represent material and non-material commodities of value [284]. Employing these concepts, the notion of a functional relationship[12] is defined to exist between any two elements in the network of actors, activities and resources described above, if the proper functioning[13] of one element depends upon, or is influenced by, the activity of another (*e.g.* an actor requiring a certain resource at a certain point in time). The network of such functional relationships contained in real-world systems, called the *functional network*, is defined to constitute the very agency by which the experience of effect propagates throughout a system to impact certain entities contained therein. The component of the functional network characterised by formal and informal interactions is called the *formal functional network*, and the *informal functional network*, respectively.

The practical selection of stakeholders

Investigating the notions of power and effect in the preceding section, the introduction of what was called a "functional network" proves to be of importance to the process of stakeholder selection. Simply put, a functional network was defined to represent the agency by which the experience of effect propagates throughout a system. At first this may seem rather insignificant, but the fact that this network represents the mechanism by which the experience of effect propagates, implies that all entities potentially influenced by the outcome of an endeavour are contained therein. By definition of what it means to be effected, an entity will only be interested in an endeavour if it is, or is likely to be, effected thereby. Based on this property, and the fact that all entities that may potentially be affected by an endeavour exist in the functional network related to it, it follows that all entities relevant to an endeavour exist in the functional network

[10] For the sake of clarity, note that outcomes formally pursued by an entity essentially represent its documented purpose within the system in which it exists. Take, for example, an entity working as a police officer. The formal purpose of a police officer, and thus one of the formal outcomes pursued by the entity working as a police officer, is to maintain order in society.

[11] In the context of this dissertation, no distinction is made between the term actor and the term entity.

[12] The existence of a functional relationship is considered formal if the reality thereof is formally documented, or informal, if the reality thereof is not formally documented.

[13] The proper functioning of an entity refers to its ability to realise the formal and informal outcomes it pursues.

related to that endeavour. This is rather an opportune property as it provides an initial filter by which to discern relevant from irrelevant entities. Notably, however, the delineation of the functional network related to an endeavour will only bring an analyst so far — all the entities relevant to an endeavour may be contained in the functional network related to it, but that does not mean that all the entities contained therein are relevant. Consequently, it is necessary to combine the notion of a functional network with some sort of guided search, which is where the dimensions proposed to govern the identification of relevant stakeholders are of service.

Power and effect are the dimensions defined to govern the selection of stakeholders. It is thus only appropriate that these dimensions be employed to guide an analyst's search within the functional network relevant to it. While designing a method encapsulating such a guided search, the following dynamic was noted to result from the definition of what it means to be affected: No matter the measure of an entity's power, unless it is actually affected by the existence of a problematic situation related to an endeavour, it remains irrelevant to the resolution thereof. In contrast to this dynamic, a lack of power fails to render the affected irrelevant[14]. Because of this dissimilar quality, the notion of effect is employed as a preliminary constraint on the functional network under investigation. More specifically, the following cyclic method[15] is proposed for facilitating the identification of potential stakeholders by means of the functional network in which they exist:

1. Formulate an initial description of the problem being addressed and the activities that surround it in as much detail as possible at this time.

2. Based on this description, identify and label the core features of the problem.

3. Using the initial description and list of features as a point of reference, construct the functional network related to it in order to identify affected entities.

4. Once satisfied that all affected entities have been identified, based upon the current description of the problem, proceed to determine which of these entities possess power relevant to the problem at hand.

5. In view of both the power possessed and the effect experienced by entities, describe the extent to which, and the manner in which, entities are relevant to the problem at hand.

There are a number of technicalities related to the execution of the above methodology that have to be addressed. First, note that although the proposed methodology is designed to elicit relevant stakeholders in both the formal and the informal components of a functional network, it is likely that the vast majority of relevant stakeholders reside in its formal component. To elucidate, consider the primary purpose of any formally defined system — the management of elements pertaining to its purpose [55]. Over time, systems develop to include entities relevant to their designed purpose. It is, therefore, expected that most of the stakeholders relevant to an endeavour will be identified by exploring the formal component of the functional network[16] related to it. As such, when attempting to identify the affected, it is suggested that analysts first attempt to construct the formal functional network, and only thereafter, the informal functional network.

[14] The refusal to dismiss affected entities because their lack of power is partially inspired by the critical systems movement — a movement that built its foundations upon the necessity of considering entities regardless of their ability to influence the system in which they exist.

[15] An analyst may iterate through the steps of the method as many times as necessary.

[16] This by no means implies that an analyst need not also explore its informal component.

The second and final technicality concerns a detail in the description of Phase 4 of the methodology that has not been addressed adequately. That is, when is an entity's power considered *relevant*? There are essentially two scenarios that qualify the relevance of an entity's power. First, if an entity's power is of significant utility toward resolving the issue under consideration or managing the activities that surround it, and secondly, if an entity's power grants it the capacity to inhibit the implementation of a potential solution to the problem considered. Importantly, when investigating the relevance of an entity's power, instead of investigating the rather ambiguous notion of power in general, the sources of power defined in the preceding section titled *The notion of power* should be utilised in order to ground the analyst's enquiry. Accordingly, the following pair of questions, collectively constituting a power analysis checklist, is proposed to facilitate the analysis of an entity's power:

1. Does the entity possess structural, customary, status, necessity or incapacitation power that is of utility toward resolving the problem considered or managing the activities that surround it?

2. Does the entity possess the structural, customary, status, necessity or incapacitation power to reject a potential solution proposed to resolve the problem considered?

In conclusion, during the review of the stakeholder selection methodologies that exist in the literature, it was noted that several of these methodologies incorporate a dimension addressing the actual level of an entity's interest in an endeavour toward determining its inclusion as stakeholder. This is, however, somewhat surprising as it is rather obvious that an entity with no interest in the outcome of an endeavour will not be included as a stakeholder. The intent of the method discussed above is not to select stakeholders on behalf of an analyst, it is rather to aid an analyst in identifying potential stakeholders. The actual selection of stakeholders remains the responsibility of the analyst. Nonetheless, it is claimed that the most relevant stakeholders may be identified by employing the guidelines introduced above.

The notion of representation

Before closing the discussion on stakeholder selection it is necessary to touch on the notion of *representation* [458] — an idea encapsulating the possibility of minimising the number of entities included in an endeavour by aggregating their involvement to an entity who is representative of the greater collective. One must act with strategic caution when minimising the number of entities within an endeavour through representation, for one must ensure that the selected entity is appropriately representative. More specifically, one should take care not to limit the information within an endeavour by employing the notion of representation carelessly. To illustrate this point, note that if entities within the original system are connected by a functional network, they may also be connected by a formal power structure (*i.e.* a formal layout of the power relationships among entities within a system). In such a case there may exist entities on a higher system level than the level on which a problematic situation of interest exists, who possess the formal decision making power to "solve the problem" without the involvement of those entities on the lower level (thereby representing them). But there is risk involved in this strategy — the measure to which the formulation of a problem is a true representation of the underlying real-world problem is critically influenced by the entities one includes within an endeavour.

As systems philosophy holds [163], developing a rich, exceedingly objective picture of a problem context greatly depends upon the degree to which different, uniquely subjective perspectives are combined. Consequently, limiting the number of stakeholders may polarise what is understood

about a problem and subsequently bias the solutions uncovered toward its resolution. In the author's view there are only two reasons that would justify such a limiting strategy, the first of which concerns the value of time. In general, the more stakeholders one includes within an endeavour, the longer it will take to reach a state of agreement among stakeholders. Consequently, one is forced at times to balance the value of developing a rich understanding of the problem context that may result in uncovering *better* solutions with the value of time spent on the problem. The second reason, partially inspired by the emancipatory priority of the critical systems movement, addresses the case in which a joint body of entities request representation by a third party in order to supplement their lack of experience in the matters of an endeavour. These cases do not ignore the possibility that if entities are connected by a formal power structure, final decision making power resides in the entity who possesses ultimate power. What is argued is that there may be value in including certain 'less powerful' entities in the problem resolution context, even if it is possible to replace them by moving to a higher level in the original system.

One final point is made in conclusion. The process of identifying relevant stakeholders has been described as the first step within an endeavour, one that must be completed before one can continue with the remainder of the endeavour. It is important to note, however, that the identification of relevant stakeholders should not be terminated after the initial stages of the endeavour. It may be the primary focus during the initial stages of an endeavour, but that does not mean that the process should be terminated as the endeavour progresses — however undesirable it may be, it is always possible that excluded, but relevant, stakeholders may be discovered during the later stages of an endeavour.

9.2.2 A selection of structuring methods

After completing the preceding phase of the proposed methodology, the purpose of the remaining phase, the structuring of the problem considered, is twofold: (1) To develop a rich understanding of the problem at hand, and (2) to decide what to do about it. In other words, it is necessary to develop a formal formulation of the problem under consideration (that accurately captures the underlying real-world problem), and to elicit the objectives that may be pursued toward its resolution. Based upon the review of the literature related to problem structuring methodologies, it is evident that a considerable amount of work has been done by past researchers to develop methods capable of facilitating this task. Notably, however, no particular method has as of yet been deemed best suited to the structuring of *every* problem type. It would seem that the distinctness of different problem structuring methodologies render them uniquely equipped to structure different types of problems. In view of the advance state of progress in the field of problem structuring, no attempt is made to reinvent the wheel. Instead, a brief analysis of a select number of problem structuring methods from the literature is presented in order to support their case-specific selection.

During the initial stages of the research reported here, the methods listed in Table 9.1 were identified and considered for inclusion within the framework proposed in this dissertation. After an in-depth investigation, however, several of the methods presented in Table 9.1 were found to be ill-suited. More specifically, the future workshop [430], drama theory [37], and scenario planning [358] methodologies are excluded from further consideration because the scope of their application is too limited. Churchman's social system design is also excluded, but this time because it is too general, or rather, it is more accurately described as a collection of philosophical meanderings rather than a systematic methodology for the structuring of problems [234]. Finally, the total systems intervention of Flood is excluded because, like the framework proposed in this dissertation, it is a *meta-methodology* (*i.e.* a methodology supporting the selection of several

methods) [163]. After filtering out the aforementioned ill-suited methods, a total of six methods remain: SSM, SODA, SCA, SAST, CSH and Interactive planning (the reader is referred to Chapter 6 for a review on the working of these methods).

TABLE 9.1: *A collection of the structuring methodologies considered for use in the framework proposed in this dissertation.*

Structuring methodology	Authors
Soft systems methodology (SSM)	Peter Checkland [102]
Strategic options development and analysis (SODA)	Ackerman *et al.* [3]
Strategic choice approach (SCA)	Friend *et al.* [178]
Strategic assumption surfacing and testing (SAST)	Mason *et al.* [307]
Critical systems heuristics (CSH)	Ulrich [458]
Social systems sciences (SSS)	Ackoff [8]
Total systems intervention (TSI)	Flood and Jackson [164]
Social systems design	Churchman [114]
Future workshop	Jungk and Mullert [256]
Drama theory	Howard [224]
Scenario planning	Kahn [17]

9.2.3 The theoretical foundation of selected structuring methods

All too often, the linkage of method to task is made not according to how well the design of a method complements the complexity of a task, but according to an analyst's history of method use. As such, the purpose of this section is to describe the sociological paradigms to which the selected methods in Table 9.1 belong, as well as the theoretical assumptions they incorporate, in order to elicit their intended purpose. As the author does not have extensive experience in the use of these methods, the exposition relies heavily upon the academic literature.

The success or failure of a comparison of multiple methods typically rests upon the selected framework of comparison — in order to succeed in this exposition, the selected framework must suit both its context and purpose. As this review is aimed at both the theoretical foundations according to which methods operate, and their practical applicability in dissimilar contexts, the following strategy is adopted. The relevant methods of Table 9.1 are first classified according to a framework proposed by Burrel and Morgan [82] toward eliciting the sociological paradigms upon which they are based and, in the next section, the practical pitfalls associated with methods that belong to certain paradigms are discussed in view of the problem typologies reviewed in Chapter 7.

In the oft cited book *Sociological paradigms and organisational analysis: Elements of the sociology of corporate life*, Burrel and Morgan [82] argued that the sociological paradigms upon which all theories of organisational science are based can be distinguished according to two dimensions: The philosophy of science they employ, and the theory of society upon which they are based. To elucidate, a particular philosophy of science is distinguished according to whether the research conducted therein is assumed *objective*, or *subjective*[17] in kind. And similarly, alternative the-

[17] In a rather significant contribution titled *Crossroads — Describing differences in approaches to organisation science: Rethinking Burrell and Morgan and their legacy*, Deetz [125] criticised the distinction between sociological paradigms along a subjective-objective dimension. He argued that such a distinction fundamentally neglects the subjective element in any approach to organisational science. Although the validity of this argument is acknowledged — it is unwise to label any form of research as entirely objective — the clarity of definitions provided by Burrel and Morgan [82] with respect to the typological dimensions they employ (including the

ories of society are distinguished according to whether they emphasise a mandate of *regulation* or *radical change*. In order to further distinguish between these dimensions, the prototypical assumptions accompanying their adoption are described below.

If an objective view of social science is employed, social reality is considered to have an objective, external existence to the individuals that function therein, human behaviour is considered deterministic[18], and quantitative tests are the preferred method for acquiring knowledge. Conversely, a subjective view of social science likens social reality to the collective aggregate of subjective consciousness, values the unique perspective of relevant entities, assumes that humans possess free will, and prefers entering the context of what is studied toward gaining detailed knowledge [82]. If a mandate of regulation is prioritised, society is essentially viewed as being consensual, and methods are designed to investigate the mechanisms by which social order (the *status quo*) is maintained. Conversely, if a mandate of radical change is adopted, societies are considered driven by the exploitation of dominated groups, and thus the consensual state thereof is simply deemed a product of that domination [82]. As such, the *status quo* is not considered interesting; instead, methods are designed to elicit the dynamics that govern radical change.

By combining the above-mentioned dimensions, four unique sociological paradigms result, namely the *radical humanist* paradigm, the *radical structuralist* paradigm, the *interpretive* paradigm and the *functionalist* paradigm, as depicted in Figure 9.4 [82]. Before categorising the structuring methods mentioned above according to these paradigms, it is beneficial to further describe them.

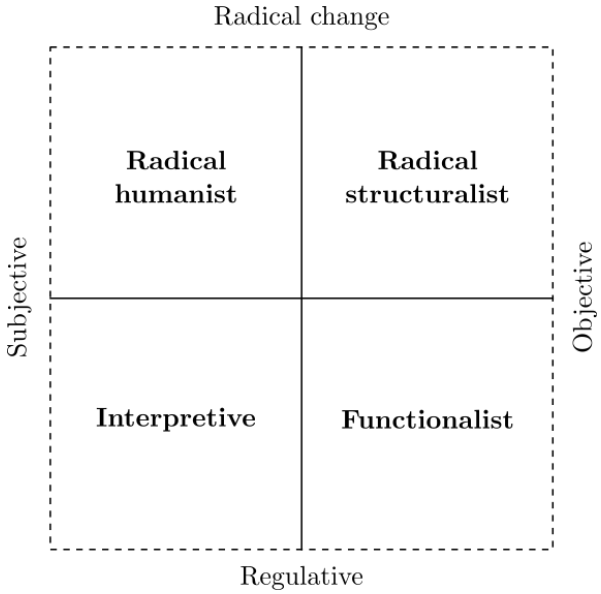


FIGURE 9.4: *Four sociological paradigms for analysing the methods of systems theory [82].*

The radical humanist paradigm, depicted in the top-left segment of the framework in Figure 9.4, is generally characterised by its dedication to sociological change from a subjective point of view. One of the most basic notions underlying this paradigm is that the *consciousness*[19] of a human entity is dominated by the ideological superstructures (*i.e.* the structure of society) in which he or she exists, and that these structures segregate that human entity from its *true consciousness*.

subjective-objective dimension) remains of great utility toward distinguishing between the sociological paradigms according to which the methods relevant to the structuring of problems operate.

[18] Determined by external circumstances.

[19] Burrel and Morgan [82] neglected to formally define what they meant by the term *consciousness*, but according to the author's interpretation, this term represents the set of beliefs, attitudes and behaviours associated with a certain individual.

The major concern of theorists who operate in this paradigm is to release human consciousness from the constraints placed thereupon by the existing social arrangements of society — a process that takes place when individuals rid themselves of the false consciousness with which they have been indoctrinated [331]. In other words, a problem solving methodology situated in the radical humanist paradigm operates from the assumption that the real world is designed in a manner that results in the oppression of one group by another, that the only manner in which to liberate the oppressed is to reveal and communicate the sources of their oppression to them, and then to facilitate radical change by facilitating the design of alternative social arrangements. In harmony with the subjective nature of this paradigm, the preferred way of learning about real-world systems is to enter their contexts and interact with the entities contained therein [234].

The radical structuralist paradigm, depicted in the top-right segment of the framework in Figure 9.4, is generally characterised by its mandate for radical change from an objective point of view. Whilst sharing an approach to science that exhibits many similarities with that of the functionalist paradigm discussed below, it does so for the purpose of fundamentally different ends. In like manner to the radical humanist paradigm, the radical structuralist paradigm functions according to the assumption that the world's design results in the oppression of certain groups, and thus dedicates itself to the liberation of the oppressed. Interestingly, however, unlike the radical humanist paradigm, it believes that the catalyst for radical change is not the enlightenment of the oppressed, but rather, that it, like the oppression of certain groups, is a natural occurring phenomenon resulting from the structure of society[20] [82]. As such, theorists who function within this paradigm study the network of elements that constitute the structure of societies around the world in an attempt to provide explanations for the manner in which this structure participates in the oppression and the liberation of human entities. As is characteristic of the objective nature of the radical structuralist paradigm, a belief in the development of quantitative models, in which human behaviour is determined by contextual elements, dominates [234].

The interpretive paradigm, depicted in the bottom-left segment of the framework in Figure 9.4, is generally characterised by a belief in the subjective sociology of regulation. In its approach to social science, the social structures of the world are considered a product of the entities contained therein and, as such, it is assumed that the only manner in which to gain insight into the nature of these structures, and the systems that constitute them, is by investigating the subjective perspectives of the free-willed human entities that participate in their existence [331]. Dedicated to this pursuit of understanding the world as it is, theorists of the interpretive paradigm seek insight from the realm of individual consciousness and subjectivity, and prefer to enter the context of the studied phenomena, essentially becoming participants in its working. The interpretive paradigm employs the assumption that the systems governing human affairs are cohesive, ordered and integrated. Because of this assumption, the problems of domination and fundamental conflict play no part in the theoretical framework upon which the interpretive paradigm is based [82].

The functionalist paradigm, depicted in the bottom-right segment of the framework in Figure 9.4, is generally characterised by a sociology of regulation, approached from an objective point of view. In the history of science, this paradigm has certainly been dominant in the study of organisations. The general purpose of the functionalist paradigm is to understand the manner in which societal order and stability exist and persist, with the end goal of improving one's ability to predict and control the behaviour of the related system. A theorist of the functionalist

[20] Common to most theorists in this paradigm is the view that contemporary society is characterised by fundamental conflicts that cause radical change in the form of either the domination of the free, or the liberation of the oppressed through both political and economic crises [82].

paradigm would argue that such understanding results when one is able to identify regularities in the behaviour of a system, and the sub-systems contained therein [234]. More often than not, this kind of investigation entails the development of analytical models in which human beings present no more problems than other components. In its overall approach, the functionalist paradigm is highly pragmatic in orientation, and dedicated to the pursuit of understanding society in a way that generates knowledge of practical utility. Essentially, it prioritises understanding order and stability in society, and how these states may be maintained [82].

Before categorising the selected structuring methods according to the paradigms defined above, there are two issues that have to be addressed. First, a number of researchers have concluded that problem structuring methods cannot be classified according to the dimensions suggested by Burrel and Morgan [82]. Essentially, these researchers argue that the practical use of structuring methods may contain both *objective* and *subjective* elements, and furthermore, that the use of structuring methods may further either a sociology of *regulation*, or *radical change* [108]. The author agrees with these arguments wholeheartedly. Note, however, that the subsequent partitioning of methods is not in respect of the practical effect of their adoption, but rather the (often subconscious) assumptions upon which their designs are based. To that end, the adoption of the framework proposed by Burrel and Morgan [82] is appropriate. The second issue is of a more theoretical nature. During the past few decades, a growing number of researchers have concluded that a belief in the classification of system methods according to incommensurable paradigms severely limits the potential for theoretical development in systems theory. The author also supports this view wholeheartedly. The subsequent method categorisation should therefore not be interpreted differently. The purpose thereof is purely to elicit the (often subconscious) assumptions upon which the design of a method is based.

Now that these issues have been addressed, consider the linkage of method-to-paradigm presented in Table 9.2. As may be seen in the table, only the CSH method does not belong to the interpretive paradigm.

TABLE 9.2: *A list of the sociological paradigms to which the structuring methodologies considered for use within the framework proposed in this dissertation belong.*

Structuring methodology	Sociological paradigm
Soft systems methodology (SSM)	Interpretive [234]
Strategic options development and analysis (SODA)	Interpretive [234]
Strategic choice approach (SCA)	Interpretive [234]
Strategic assumption surfacing and testing (SAST)	Interpretive [233]
Critical systems heuristics (CSH)	Radical humanist [232]
Social systems sciences (SSS)	Interpretive [234]

9.2.4 The practical pitfalls of paradigm-based methods

Several authors have attempted to elicit the potential pitfalls related to methods that adopt certain sociological paradigms. The work of these authors is employed to present the reader with a guideline for the risks associated with the adoption of certain structuring methods according to the sociological paradigms upon which they are based. The subsequent discussion is not intended to discourage the use of structuring methods. It is rather meant to inform practitioners of the ways in which the performance of the various methods, based upon certain sociological paradigms, are susceptible to unwanted contextual influences so that their application can be managed accordingly. As the methods in Table 9.2 belong to only the interpretive and the

radical humanist paradigms, the subsequent discussion neglects the radical structuralist and the functionalist paradigms.

The interpretive paradigm

In their purest form, methods that belong to the interpretive paradigm are designed to facilitate the process of identifying a response to problematic situations which is theoretically desirable, practically feasible, historically appropriate, culturally viable and politically possible. Although interpretive methods have certainly established their utility in this regard, there are several critiques thereof that should be noted. The main critique of methods based upon the interpretive paradigm is not so much a critique on the methodological processes they employ, but rather a critique on the manner in which they are employed. More specifically, critics argue that practitioners fail to recognise the limited applicability of interpretive methods [233]. The interpretive paradigm is based on a consensus world-view in which conflicts of interest can always be settled through the process of civil debate. Critics argue, however, and rightly so, that many problematic instances do not exhibit this characteristic, and that the ignorant application of interpretive methods in a coercive context, as described in the final typology of §7.4, will simply benefit the powerful [232, 320]. Essentially, interpretive methods are not designed with the inherent ability to accommodate the dynamics of a coercive context and their application should therefore be managed accordingly.

A further critique of interpretive methods is their reliance upon the idealised interaction of entities. One of the core principles of the interpretive paradigm is that the discovery of appropriate solutions is significantly aided by a rich understanding of the related problematic situation which, in turn, is the product of considering the unique perspective of competing entities. Despite the truth of this hypothesis, practitioners of interpretive methods typically fail to recognise that certain entities may be inhibited by internal or external factors from speaking their minds freely (*e.g.* in a hierarchical setting) [234]. It is rather obvious that, if left unchecked, powerful entities may influence and dominate the methodological processes of methods based upon the interpretive paradigm and, as such, their application should be managed accordingly. More specifically, their adoption may be better suited to problem instances that are unitary, or pluralist in kind, as described in the final typology of §7.4.

The radical humanist paradigm

Designed to address societal issues of oppression, structuring methods based upon the radical humanist paradigm embody a refusal to accept the inequalities of status, wealth, power and authority in society. This is no doubt a praiseworthy mandate, but the following question arises: To what extent are the methods of this paradigm able to deliver?

The main critique levelled at methods of the radical humanist paradigm is rather an interesting one. Radical humanist methods employ the assumption that oppressed human entities will attempt to liberate themselves if they are able to recognise the reality and the source of their oppression. Although this hypothesis is generally accepted, the assumption that such insight is readily available is not. Critics employ the work of philosophers and sociologists that are of a *materialist* persuasion (*e.g.* Marx and theorists of the Frankfurt School) to illustrate that values and beliefs develop, that they are related to the political and economic aspects of the time, and that the very structure of society determines that certain ideologies dominate others at particular times [233]. Essentially, critics of the radical humanist paradigm argue that it is exceedingly difficult for either entities participating in an endeavour, who have typically become so accustomed to their circumstances that they fail to register the need to be liberated, or expert

analysts, who are predominantly at the mercy of conditioned participants for insight, to identify and make salient the sources of oppression.

Almost all methods of science contain weaknesses, and in view of the above critique it is clear that the problem structuring methods considered in this section are no different. Despite these weaknesses, the analysis conducted in this section should support the selection and use of the structuring methods in Table 9.2 in a manner that is beneficial. In conclusion, it is wise always to keep in mind that no mere methodology, which is but words on paper, can guarantee the success of an enterprise. How successful a particular methodology turns out to be will depend upon several factors including, but not limited to, the practitioner thereof and the particular situation in which it is employed.

9.3 A decision-making checklist

It is evident from the literature review of Chapters 3 and 4 that the mechanisms of human decision making remain vulnerable to irrational influences. As of yet, no attempt has been made to address the potentially adverse effect of such influences during the application of the proposed methodology. Toward doing so, a debiasing paradigm proposed by Fischoff [159], titled *A mismatch between judge and task*, is of significant utility. In fact, the principles that govern this paradigm are adopted as the conceptual framework upon which the debiasing framework presented subsequently is based. The following creed illustrates the perspective adopted in the aforementioned paradigm: An instance of poor decision making results not from the inherent ability of decision makers, nor from the structure of a decision making task, but from the mismatch of these elements[21] (see §5.2 for more information).

As reviewed in §5.2, Fischoff [159] proposed five general debiasing strategies for the mitigation of biases induced by the mismatch of judge and task. Of these, the following three are drawn upon during the design of the debiasing framework presented subsequently: (1) Force participants to express what they know explicitly, instead of letting their thinking remain "in the head", (2) encourage participants to search for discrepant evidence, rather than only focusing on details that corroborate a preferred answer, and (3) get participants to consider alternative situations related to the one under consideration toward better understand the situation at hand. Before introducing the debiasing framework inspired by the above strategies, it is necessary to explicate the two types of error, which seem to dominate the irrational influence of human decision making, that it is designed to address.

The reader is reminded of the notion of cognitive biases which was the predominant focus of Chapter 3. Cognitive biases were defined as the potential for error inherent in the adoption of judgement heuristics. In a sense they represent the case in which the cognitive mechanisms of the mind fail to respond appropriately to a set of decision options presented, thus resulting in poor decision making. This potential error in the mechanisms of the mind, subsequently called the *cognitive error*, represents the first type of error addressed by the proposed decision-making checklist. The second type of error is somewhat more difficult to describe.

There is a subtlety in the social mechanisms of the self that should be noted. On the one hand, the social mechanisms of the self act as the foundation from which individuals interact with any form of social stimuli. On the other hand, built upon this foundation, every stimulus elicits a response that individuals are able to employ as a guide for subsequent decision making. The manner in which the use of these responses may result in poor decision making is contained in

[21] Despite proposing alternative paradigms, Fischoff [159] suggested that a mismatch between judge and task is likely the most appropriate debiasing paradigm, given the working of the human psyche.

TABLE 9.3: *A decision-making checklist.*

Dynamics addressed	Architecture
Anchoring bias & Affective failures	The tendency to satisfice, even when significantly better solutions may be attained with little further effort.
	Reconsider the selected decision options. In what way can these courses of action be further improved? Are these improvements worth pursuing?
Affective failures	The tendency to downplay the consequences of a decision option if those consequences are not personally experienced, or if they are realised over an extended period of time.
	Reconsider the selected decision options. What are the long-term consequences of the various decision options? Describe the worst-case scenario. Are these outcomes acceptable?
Optimistic bias, Illusion of control & Misconceptions of chance	The tendency to be overly optimistic about the future and one's ability to realise outcomes pursued.
	Reconsider the selected decision options. Make a list of the risks and uncertainties associated with the respective options. Are these risks appropriate? What steps can be taken to render the future less volatile?
Illusion of validity, Insensitivity to predictive accuracy & Confirmation bias	The tendency to prematurely accept the validity of sourced information if that information coincides with one's preconceived beliefs, while simultaneously neglecting information that contradicts those beliefs.
	Reconsider the information upon which the selected decision options are based. Is the validity of this information apparent? Formally state/restate the information that contradicts the appropriateness of the selected decision options. On what grounds is this information rejected?
Groupthink, Abilene paradox & Shared information bias	The tendency to withhold information believed to be contrary to the opinion held by the dominant players in a context, particularly when that information is possessed only by a single individual.
	Reconsider the various decision options. In an anonymous fashion, request that participants describe any issue related to the realisation of the various decision options that have not yet been addressed sufficiently.
Group polarisation & Discontinuity effect	The tendency whereby the point of view adopted by an entity consisting of several members is more extreme than that of an entity consisting of a single member.
	If reaching consensus proves to be a persistent challenge, and the above tendency is suspected, an analyst may request that all entities select a single representative of their cause, in order to organise an "executive meeting" with only the selected individuals.

the case of the cognitive error described above. But the manner in which changes in the very social state of the self, upon which these responses are often based, may result in poor decision making has yet to be noted. In such cases, the foundation from which decisions are being made is temporarily misaligned with respect to the objectives an entity would have pursued if he or she had not been influenced by the social dynamics of the environment. The misalignment of objectives in this manner, because of a temporary change in the social mechanisms of the self, subsequently called the *social error*, represents the second type of error the decision-making checklist is designed to address[22].

Now that all technicalities have been addressed, a decision-making checklist, designed for consideration at strategic junctures during an endeavour, is proposed[23]. As depicted in Table 9.3, the decision-making checklist contains several subsections, each addressing a unique set of dynamics (cognitive and/or social), grouped according to the similarity of the practical consequences in which they result. Each of these subsections employ the following structure: First, the set of dynamics being addressed is listed in the column on the left, then, in the column on the right, a short description of the manner in which the listed dynamics may manifest themselves in the context of real-world problem solving, along with an enquiry designed to mitigate the influence of such dynamics, is provided in a colour-coded fashion (framed in purple and green, respectively). For further information on any of the cognitive and social dynamics addressed in the decision-making checklist, the reader is referred to Chapters 3 and 4, respectively.

In conclusion, the decision-making checklist in Table 9.3 is designed to be employed in combination with the structuring methods discussed in §9.2.2–9.2.4 toward addressing the cognitive and social dynamics of human decision making that are typically neglected therein.

9.4 Chapter summary

The capacity to comprehend and then formulate a representation of the real-world that can generally be understood and agreed upon is arguably one of the first steps toward responding appropriately to the difficulties contained therein. As such, this chapter was dedicated to the exposition of a meta-methodology designed to facilitate such a formulation. The methodology was first introduced in §9.1, and this was followed by a two-part discussion on its constituent components. More specifically, an exposition of the various phases involved in the process of problem formulation was presented in §9.2, and this was followed in §9.3 by a description of a decision-making checklist designed to mitigate the irrational influence that exists in the mechanisms of human decision making.

[22] Note that not every dynamic described in the literature review of Chapters 3 and 4 is addressed in the decision-making checklist. The reason for this is simply that certain cognitive and social dynamics are so inherent to human nature that their influence is nearly impossible to mitigate (*e.g.* the ingroup-outgroup bias or the retrievability bias).

[23] The exact point in time is left to the analyst's discretion.

CHAPTER 10

A theoretical case study

Contents

In accordance with the research methodology adopted in the dissertation, this chapter is dedicated to an identification of shortcomings in the proposed meta-methodology during its hypothetical application to a historical, real-world endeavour (*i.e.* an endeavour that transpired as a result of a movement, in the South African eduction system, known by the phrase *fees must fall*). As this phase in the evaluation of the meta-methodology is hypothetical in nature, only the method proposed to facilitate the selection of stakeholders is considered. The application of this method is, furthermore, presented only to the extent that the description and amendments aimed at mitigating the identified shortcomings are facilitated. The purpose of this chapter is thus not to provide an entirely accurate or holistic view of the application of the evaluated method, or the fees must fall movement.

Notably, this chapter also contains the introduction of an additional phase in the proposed meta-methodology designed to address a potentially detrimental phenomenon uncovered during the hypothetical application of the methodology.

The reader is provided with a brief overview of the fees must fall movement in §10.1. Thereafter, the shortcomings of the method proposed to facilitate the selection of stakeholders, as identified during its hypothetical application in the context of the aforementioned movement, are made salient and addressed in §10.2. In §10.3, both the technicalities that underlie the design of the additional methodological phase and the constituent method is introduced. The chapter finally comes to a close with a short summary in §10.4.

10.1 The fees must fall movement

At the time, the declaration of a democratic South Africa in the year 1994 brought with it the promise of many good things, including a future in which the joys of education would be available

to all. Consider, for example, the following excerpt from the Constitution of the Republic of South Africa, adopted in 1996 [119]:

> "Everyone has the right (a) to a basic education, including adult basic education; and (b) to further education, which the state, through reasonable measures, must make progressively available and accessible" [119].

Unfortunately, after more than two decades, rights of the kind described above are still nothing more than empty promises [269]. In October 2015, the persistent failure of the South African government to deliver on such promises finally culminated in the emergence of a movement dedicated to the liberation of education from the bonds of historical oppression, known by the phrase *fees must fall* [290]. As a result of this movement, several primary role-players, including former President Jacob Zuma, were finally forced to address the dire state of education in South Africa. The remainder of this section is dedicated to an overview of the fees must fall movement in an attempt to provide the reader with sufficient insight into its development.

The exact date of the fees must fall movement's inception is generally dated to the 14th of October 2015, a day on which students of the University of the Witwatersrand took to the streets in protest of a proposed 10.5% increase[1] in fees for the 2016 academic year, resulting in the temporary closure of the University of the Witwatersrand [277]. The movement's primary, initial goal was that academic fees remain unchanged for 2016 and, unsurprisingly, it did not take long for protests in support of this mandate to spread. By the 20th of October, the call "fees must fall" echoed around institutions of higher-education throughout the country, with additional protests taking place at Rhodes University, the University of Cape Town, the University of Fort Hare, Stellenbosch University and the Cape Peninsula University of Technology [277]. Many of these universities responded to protests by employing the service of the police and that of private security companies to discourage the aspirations of potential protesters[2], thereby attempting to dampen the flame of revival that threatened to engulf the South African education system. Interestingly, several analysts argue that the employment of police in dealing with protesters increased the violent nature of the fees must fall movement sevenfold [277, 290].

On the 21st of October, protests continued to escalate with students from the University of Cape Town, the Cape Peninsula University of Technology and Stellenbosch University collectively marching to the houses of parliament in Cape Town, where former President Zuma and former Minister of Higher Education and Training, Blade Nzimande, were supposed to address them in respect of a memorandum of demands submitted several days prior [355]. When protesters finally reached parliament, however, neither President Zuma nor Minister Nzimande were there[3]. To remedy the situation, in his capacity as acting Minister of Higher Education and Training, Thula Nxesi attempted to read a response to protesters, but his address was disrupted by protesters who were upset that neither President Zuma nor Minister Nzimande were present [355]. In their frustration protesters turned to violence, hurling burning cardboard boxes (among other objects) at police. After several moments of violent exchange, riot police dispersed the crowd of protesters using stun grenades, tasers, tear gas and batons [290]. The following day, protesters in Johannesburg furthered the cause by marching to Luthuli House in a similar fashion, handing over an additional memorandum of demands to the *African National*

[1] The proposed fee increase was predominantly the result of several unavoidable increases in University expenses, such as a 7% increase in academic staff salaries and increases in foreign-sourced academic material due to the diminishing value of the South African Rand [372].

[2] In addition to these measure, the University of Cape Town selected to obtain court interdicts that resulted in the arrest of several protesting students [290].

[3] According to certain sources, President Zuma was evacuated because several protesting students had gained access to the parliament buildings [290].

Congress (ANC) secretary general, Gwede Mantashe [298]. On the 23rd of October, all the events described above, and several not mentioned here, finally resulted in former President Zuma announcing on live television that the demand of no increase in academic fees for 2016 would be met [22].

At this juncture, the entities participating in the protests surrounding the fees must fall movement were no longer limited only to university students. From an outsider's perspective, protesters may have seemed of one mind, but under the guise of the fees must fall movement there were, in fact, several groups of protesters, each with there own, at times overlapping agendas (*e.g.* university staff protesting poor retirement plans, and several political groups with agendas unknown) [431]. Consequently, although several further protests took place in the weeks that followed then President Zuma's announcement, many of these were no longer dominated by the issue of academic fees. In fact, protests dedicated to the cry of fees must fall slowly died down during this period [431]. This more peaceful period, however, turned out to be like the quiet eye of a hurricane — it did not last.

Compared to the detailed documentation of the 2015 fees must fall movement, the events surrounding the movement in 2016 are significantly less well documented. The movement's revival seems to have occurred in response to a recommendation made by the Council of Higher Education to the Minister of Higher Education and Training of the time, Blade Nzimande, with respect to the planned increase in fees for the 2017 academic year [9]. The council considered another 0% increase in fees to be unsustainable and, as such, recommended that Minister Nzimande allow universities to employ an inflation-related increase for 2017, adding that it may be wise to leave the final decision to the management of each university [451]. Almost instantaneously, this recommendation resulted in the revival of the fees must fall campaign under the hashtag #FeesMustFall2016. Minister Mantashe attempted to calm students nationwide by pointing out that no final decisions had been made and that the announcement was merely a preliminary recommendation, but it was too little too late as protests commenced nationwide [9].

This time, the primary goal of the movement was not merely a 0% increase in fees, but rather, that the system of higher education in South Africa be "decolonised" and made completely free of charge[4]. Significantly, compared to 2015, the 2016 wave of student protests were of a slightly different kind: In 2016, considerable emphasis was placed on the fact that the fees must fall movement should remain leaderless[5] as the students participating therein apparently despised the idea of a single individual becoming "the face" of their collective cause [343]. A further, less political difference between the movement in 2015 and in 2016 is the fact that student protests in 2016 were significantly more[6] violent [152]. Collectively, the aforementioned dynamics seem to underlie the final dissimilar quality of the #FeesMustFall2016 movement — the presence of considerable division among the protagonists thereof regarding the means by which outcomes were being pursued [343].

The distinctive qualities of the #FeesMustFall2016 movement, as described above, also seem to have been the primary catalyst of its conclusion. To elucidate, as the #FeesMustFall2016 movement progressed, the use of violence by both protesting students and the South African Police Service seemed to escalate unabated [447, 484] and, as a result, many students were no longer willing to participate in the protests surrounding the fees must fall movement (either

[4]Interestingly, a study by the Centre for the Study of Violence and Reconciliation concluded that the prioritisation of these goals (as well as several other factors) contributed to the #FeesMustFall2016 movement being less successful than the year prior.

[5]Recall that in 2015 it was clear that the *student representative council* (SRC) leader Nompendulo Mkhatshwa, and her predecessor Shaeera Kalla were leading the movement [343].

[6]One cannot help but wonder whether the excessive use of violence in the fees must fall movement of 2016 was partially the bi-product of its leaderless nature.

because of a moral conviction against the means being employed, or for fear of injury at the hands of the police). Unsurprisingly, this loss of support resulted in the #FeesMustFall2016 movement slowly coming to a halt several months after its revival.

The fees must fall movement may have ended in a rather distasteful fashion[7], but since its conclusion, it has become clear that the protagonists thereof accomplished much of what they had set out to do. In February 2018, the government finally released a statement officially confirming that, and how, it planned on aiding students financially. Accounting for all commitments made, an additional amount R57-billion was allocated to aid students from poor or working-class families with a combined total annual income of less than R350 000. The planned timeline of these allocations were as follows: R12.4-billion would be made available in 2018, R20.3-billion in 2019 and R24.3-billion in 2020.

10.2 The identification of potential stakeholders

Imagine a network of interconnected entities, all contained in, and relevant to, the working of a certain real-world system. It is rather obvious that although such entities may all be relevant to the working of the system in which they exist, if a problematic situation were to emerge therein, some entities would be more relevant to the process of addressing it than others. This nuanced nature of an entity's relevance may be rather simple to acknowledge, but the actual separation of relevant from irrelevant entities is anything but simple. As stated in §9.2.1, the identification of relevant stakeholders remains one of the significant obstacles in the road of real-world problem solving. Consequently, the notions of power possessed and effect experienced were suggested as dimensions that should govern the inclusion or exclusion of entities in an endeavour in order to support practitioners in this task. More specifically, it was argued that the key to identifying relevant stakeholders resides in an ability to describe theoretically, and recognise practically, the building blocks that comprise the notion of power, as well as the features that constitute the notion of effect. As the purpose of this section is to uncover any shortcomings of the method proposed in Chapter 9, the phases constituent of the method are summarised below (for the sake of convenience):

1. Describe the problem considered in as much detail as possible at this time.
2. Based on this description, identify and label the core features of the problem.
3. Proceed to employ these features and the preceding description as points of reference to conduct a search for entities affected by the problem considered in the formal and informal functional networks related to it.
4. Once satisfied that all affected entities have been identified, evaluate which of these actually possess power relevant to the problem at hand.
5. Finally, in view of the analyses conducted during both Phases 3 and 4, describe the extent to which, and the manner in which, entities are relevant to the problem considered.

From this summarised description, it is again evident that each progressive phase in the proposed methodology acts as a filter separating relevant from irrelevant entities. That being said, this filtering process is embedded within a method designed to be executed in a cyclic fashion, thereby attempting to inhibit the failure to include relevant entities (*i.e.* during each iteration of the

[7] Damages to educational infrastructure stood at nearly R800-million by the beginning of 2017 [446].

methodology, the execution of its constituent phases may be re-assessed and refined accordingly). To commence the evaluation of the above method, consider the subsequent summary of an hypothetical application of its first few phases.

As stated in §10.1, the inception of the fees must fall movement[8] is generally pinpointed as the day students from the University of the Witwatersrand elected to protest the increase in fees planned for the 2016 academic year. Consequently, this period is taken as the juncture at which to attempt the identification of stakeholders. More specifically, the application of the proposed methodology is evaluated for the period during which student protests were rampant at the University of the Witwatersrand, but before such protests became commonplace throughout the country. Accordingly, in fulfilment of Phase 1 of the proposed methodology, the problem addressed during this hypothetical stakeholder analysis is described as follows:

> A select portion of the student body at the University of the Witwatersrand is unable to afford the planned increase in academic fees for 2016 and have elected to exercise their right to protest in objection, demanding that fees remain unchanged.

In view of the above preliminary description of the problem considered, the following contributing features are considered (in fulfilment of Phase 2): Financial issues[9], security issues and operational issues. With a preliminary set of core features identified, the next step in the proposed method involves constructing first the formal functional network, and thereafter the informal functional network, related to each feature (in fulfilment of Phase 3). That is, the next question of interest toward identifying the affected in the system under consideration is phrased as follows: Which entities in the functional network related to the problem considered pursue formal outcomes affected thereby?

As the University of the Witwatersrand constitutes the backdrop of the problem considered, the functional network related to it is employed to facilitate the identification of the effected. For no specific reason, furthermore, the core features of the problem considered are investigated in the order listed. Consequently, the above question of interest is rephrased as follows: Which entities in the functional network related to the working of the University of the Witwatersrand pursue formal outcomes affected by the financial issues associated with the problem considered? In pursuit of answering this question, students enrolled at the University of the Witwatersrand are deemed severely affected by the issue of finance — their primary, formal purpose of attaining a tertiary qualification would be severely complicated if the proposed increase in tuition fees were to be formalised[10]. According to the propagation of effect, furthermore, all entities formally responsible for supporting students financially are also considered formally affected.

A further group of effected entities uncovered by constructing the functional network related to the question of interest are those tasked with the responsibility of managing the financial affairs of the university. Such entities include, but are not limited to, the Council of the university[11],

[8] In the opening section of Chapter 9, the meta-methodology proposed in this dissertation was described as one designed to address problems that exist in real-world systems in which entities are forced to collaborate despite pursuing conflicting outcomes. It should be rather obvious that the system of higher education in South Africa satisfies this property and, as such, is a context well-suited to an illustration of the use of the method proposed to facilitate the selection of stakeholders to an endeavour.

[9] The central role of finance in the problem considered is an expected one. As stated in §9.2.1, the primary function of any formally defined system is the management of elements pertaining to its purpose, and as the purpose of most real-world systems include some kind of financial incentive, it is only natural that the resolution of problems contained therein entail a financial dimension.

[10] As the severity of effect is determined with respect to an entity's ability to realise the outcomes it pursues, note that students less able to accommodate the proposed increase in fees are considered more severely affected thereby.

[11] The Council of the University of the Witwatersrand constitutes its highest decision making body [460].

the financial committee of the Council led by the chief financial officer, the Vice-Chancellor's team[12], the financial aid, scholarships and merit awards committee of the Senate, the financial aid office, and finally, the university's fees office [461, 459]. The next instance of effect associated with the problem considered are the security issues that surround it. Figure 10.1 depicts the state of the University of the Witwatersrand during the initial protests[13] and stands as exemplar of the potential for chaos hidden in the dynamics that surround protests [138]. By constructing the functional network induced by this issue, all entities pursuing formal outcomes in close proximity to protests, and, as a result, all entities formally responsible to defend and protect society from the kind of dangers that surround protests (*e.g.* the police force) were identified as potential stakeholders.

In addition to the issue of security, Figure 10.1 also alludes to one of the operational issues[14] caused by student protests — restricted access to the University of the Witwatersrand's campus. Access to the university's campus is fundamentally linked to the realisation of the formal outcomes pursued by the majority of entities contained in the formal functional network related to the University of the Witwatersrand. Consequently, by tracing out the functional network associated with the feature of access to university grounds, the following group of affected entities were additionally identified: Non-protesting students, professors, lecturers, administrative personnel, and general visitors to the university.

FIGURE 10.1: *A scene captured at one of the entrances to the University of the Witwatersrand campus during the protests in response to the proposed increase in fees for the 2016 academic year [348].*

The hypothetical application of the third phase of the proposed methodology resulted in the identification of several shortcomings related thereto. As it is deemed beneficial to address these issues before continuing to the next phase of the methodology, the account of its application is paused here. The first shortcoming concerns the construction of the formal functional network related to the identified features of the problem considered. More specifically, while constructing this network it was unclear at times whether or not a certain entity is able to absorb the measure of effect it experiences, or whether this effect propagates beyond the entity in question to the more general network in which it resides. This resulted in being ambiguity in whether or not certain entities form part of the functional network of interest. A particularly relevant example of this dynamic relates to the propagation of effect in the context of systems characterised by

[12]The Vice-Chancellor's team is responsible for managing the university on a daily basis.

[13]Figure 10.1 is not an accurate representation of the racial distribution of the students that partook in the fees must fall movement.

[14]Note that any operational issue resulting from an underlying financial one is neglected as entities affected by such issues were identified indirectly during the investigation of financial issues.

hierarchical levels of governance as this kind of real-world system is essentially designed to limit the propagation of effect. That is, if a problem occurs on a particular level of governance in such a system (*e.g.* provincial authority), the ripples of effect induced thereby only propagate to a higher level in the governance structure of the system (*e.g.* national authority) if the entities on the previous level are unable to address the effect-inducting problem appropriately (*i.e.* if they do not possess the power to do so). The difficulty in respect of the execution of Phase 3 is that the analysis of an entity's power takes place only during Phase 4 of the methodology and, as a result, it is difficult to ascertain whether or not the propagation of effect should be limited to a certain entity (*i.e.* a certain level of a hierarchical system) during the execution of Phase 3. To illustrate this dynamic in the context of a practical example, consider the following scenario related to the propagation of effect induced by the inception of the fees must fall movement at the University of the Witwatersrand.

Section 10.1 opened with an excerpt from the Constitution of the Republic of South Africa stating that every citizen has a right to higher education, which the government, through reasonable measures, must make progressively available and accessible. Unsurprisingly, the same document states that the governmental body responsible for ensuring such a state of affairs finds its apogee in the President of South Africa. To elucidate, the Constitution, dictates that the governance of education in South Africa is structured such that there exists a department of education in each of the country's nine provinces, and that each of these departments is responsible to the minister of higher education for the state of education in the province in which it resides. The minister of higher education is, furthermore, responsible to the President, who is responsible to the citizens of South Africa for the state of higher education in the entire country [129]. By employing their right to protest, the students of the University of the Witwatersrand essentially caused the ripples of effect produced by the financial, operational and security issues associated with the problem considered to propagate into the functional network contained in the aforementioned structure of governance dedicated to higher education in South Africa. In order to identify the affected, the question of interest is thus: How far up this governance structure did these ripples of effect go? The difficulty in answering this question during the initial execution of Phase 3 is a case in point of the aforementioned shortcoming in the current version of the proposed method. In order to address this shortcoming, the following strategy is proposed:

Amendment 1. If it is unclear whether or not the experience of effect propagates beyond a certain entity during the execution of Phase 3, an analyst should proceed by assuming that the propagation of effect is, in fact, limited to the entity in question. Thereafter, during the execution of Phase 4, the validity of this assumption should be evaluated. That is, an analyst should determine whether the entity in question does, in fact, possess the power required to limit the propagation of effect. If the assumption is found to be wanting, Phase 3 of the methodology must be repeated and the formal functional network expanded accordingly.

In the context of the aforementioned scenario, it is thus reasonable to assume, at least initially, that the experience of effect in the governmental body dedicated to higher education is limited to the Department of Education in the Province of Gauteng (the validity of this assumption is evaluated subsequently).

The hypothetical application of the proposed methodology to the context of the fees must fall movement facilitated the identification of another shortcoming in the working of Phase 3. In the current version of the proposed method, the task of identifying affected entities contained only in the informal functional network of interest was found to be excessively difficult. Evidently, the identification of relevant entities not formally documented requires more from an analyst

than simply uncovering a number of elusive, but formally documented functional relationships — it requires an analyst to envisage such interdependencies based upon incomplete information. Toward better facilitating this task, a supplementary technique is proposed. Consider the following phrase penned by Karl Marx [70]:

"History repeats itself, first as a tragedy, second as a farce" — Karl Marx.

The pages of history were found to be exceptionally useful toward providing a starting place for the identification of the informally affected. From the privileged perspective of hindsight, the author was able to employ historical examples as a source of insight toward identifying affected entities not formally documented. For example, by examining the history of tuition fees in the United Kingdom, prospective students planning on, but not yet enrolled, at the University of the Witwatersrand were identified as informally affected [373]. Accordingly, the following strategy is proposed to facilitate the development of the informal functional network related to a problem considered:

Amendment 2. Analysts are advised to draw upon the utility of history to elicit affected entities hidden from purely an analysis of real-time information. This can be achieved by reviewing historical endeavours pertaining to problems similar to the one considered, and utilising the accompanying documentation as a starting point for the identification of the informally effected.

With the aforementioned shortcomings related to the working of Phase 3 addressed, the discussion now returns to the hypothetical application of Phase 4 of the proposed methodology. As stated during the introduction of the proposed methodology in §9.2.1, the purpose of Phase 4 is to determine which of the entities deemed affected during Phase 3 actually possess power relevant to the problem at hand. To facilitate this process, a power analysis checklist, based on the two scenarios that render an entity's power relevant, was proposed. For the sake of clarity, the pair of questions contained therein are restated: (1) Does the entity possess structural, customary, status, necessity or incapacitation power that is of utility toward resolving the problem considered or managing the activities that surround it? (2) Does the entity possess the structural, customary, status, necessity or incapacitation power to reject a potential solution proposed to resolve the problem considered? To facilitate the evaluation of these questions in the context of the current hypothetical stakeholder analysis, a summary of select entities deemed affected is presented in Table 10.1. Note that the depth of the analysis surrounding dissimilar entities is varied as deemed appropriate.

During the execution of Phase 4, the formal power structure of the University of the Witwatersrand was found to be of considerable utility. This power structure revealed that a certain body of governance, called the Council of the university, possesses absolute structural authority to determine the rate of academic fees imposed upon the students of the University of the Witwatersrand [459]. In view of this reality, it may be tempting to conclude that no other entity need be included in the process of addressing the problem considered. Such a conclusion would, however, be detrimental. First, as illustrated in the power analysis checklist, there are multiple forms of power that must be considered when identifying the relevant, and furthermore, even if an entity does possess the capacity to formally select a certain course of action, which does not imply that such an entity is capable of selecting that course of action appropriately. To illustrate this dynamic which is, in fact, the motive underlying the first of the two questions that constitute the power analysis checklist, note that throughout a typical year at the

TABLE 10.1: *Entities deemed affected during the execution of Phase 3 of the initial stakeholder analysis.*

General category	Description
Students of the University of the Witwatersrand	Emphasis placed on financially struggling students
Managerial, administrative or academic personnel of the University of the Witwatersrand	Council of the university
	Financial committee of the Council
	The Vice-Chancellor's team
	The financial aid, scholarship and merit awards committee
	University financial aid office
	University fees office
	Professors, lecturers and administrative personnel
Governmental body	Gauteng Department of Education
	Gauteng Police Department
General public	Entities in close proximity to the University of the Witwatersrand

University of the Witwatersrand, its Council deals with only the strategic management of the university and consists predominantly of individuals not officially employed by the university (roughly 60%) [460]. Contrast this with another entity formally a part of the University of the Witwatersrand also deemed affected during Phase 3: The Vice-Chancellor's team, whose members are responsible for the management of the university on a daily basis [461]. Because of the unique nature of the respective roles fulfilled by these entities, it is quite possible that the members of the Vice-Chancellor's team, although structurally less powerful than the Council, are much more informed in respect of the state and working of the university on a daily basis. Moreover, given their daily interaction with members of the university, it is likely that certain members of the Vice-Chancellor's team are acquainted with key role-players in the fees must fall movement. This kind of consideration regarding information possessed or individuals known forms part of the analysis surrounding status power, and in view of the above discussion, the intended purpose of Question 1, and the importance of considering the entire spectrum of power sources deemed relevant to both Questions 1 and 2 of the power analysis checklist, is evident.

Consider the logic behind the second question proposed to guide the analysis of an entity's power: If an entity possesses the power to disrupt the practical implementation of a solution proposed to a certain problem, the only logical precaution is to involve that entity in the actual process of resolving the problem considered. In that way, the entity can employ the power it possesses, in the sense described in Chapter 13, to influence the proposed solution according to its capacity. Consider the power possessed by the student body at the University of the Witwatersrand. If it is possible to describe a certain entity as the primary driving force behind an endeavour, it implies that the entity possesses the power to reject potential solutions proposed to the problem considered[15]. Consequently, as the students of the University of the Witwatersrand undoubtedly constitute the driving force behind the pending endeavour, the question of interest is therefore not whether the student body at the University of the Witwatersrand possesses the power to reject potential solutions proposed to the problem considered, but rather where this power comes from. Without going into too much detail, it follows from the cited documentation

[15] An entity is only able to cause the organisation of an endeavour if it possesses the power to reject the current state of affairs, which implies that it also possesses the power to reject any proposed future state of affairs.

surrounding the fees must fall movement that the majority of power possessed by the students of the University of the Witwatersrand exists in the form of customary, necessity and incapacitation power.

The Gauteng Police Department is one of the few entities deemed affected in Phase 3 that exist outside the formal structure of the University of the Witwatersrand. Interestingly, during the evaluation of the power possessed by that entity, its power seemed either of considerable utility, or complete irrelevance, depending on the aspect of the problem considered. This polarity in the relevance of the Gauteng Police Department is considered the product of a shortcoming in the current version of the proposed method. This shortcoming led to the formalisation of another supplementary technique — the segregation of the description of the problem considered into the parts that constitute it. To elucidate, consider the following three-part segregation of the problem considered: (1) A certain number of students at the University of the Witwatersrand are unable to afford the proposed increase in academic fees for 2016, (2) students have elected to exercise their right to protest in objection to what they considered as excessive academic fees, and (3) protesting students are demanding that fees remain unchanged for the 2016 academic year. By segregating the problem considered in this fashion, it is clear that the Gauteng Police Department is irrelevant to every aspect of the problem considered, except with regard to Part 2 — student protests. In terms of student protests, the structural, customary, status and incapacitation power possessed by the Gauteng Police Department causes few entities to be more relevant. This exemplifies the potential value of segregating multifaceted problems toward the identification of entities that possess power of significant utility to the resolution of one such facet, but completely irrelevant to another. Accordingly, consider its formal adoption into the working of Phase 4 of the proposed method:

Amendment 3. During the execution of Phase 4, analysts are advised to segregate a problem considered into its constituent parts, thereby facilitating a more nuanced analysis of an entity's power with respect to each of these parts.

Before concluding the discussion dedicated to a review of Phase 4 of the proposed methodology, recall the assumption made during Phase 3 regarding the propagation of effect within the governmental body responsible for higher education in South Africa. Simply put, the Gauteng Department of Education was assumed to be the only entity in this sector affected by the problem considered. As stated, this kind of assumption is appropriate in the context of Phase 3, but during the execution of Phase 4, the validity thereof should be questioned. To that end, note that the Gauteng Department of Education was deemed the subject of the aforementioned assumption, not by chance selection, but because the formal outcomes it pursues were considered relevant to the resolution of the problem considered [184]. It is therefore necessary to confirm that the Gauteng Department of Education did, in fact, possess the power to realise the formal outcomes it pursues despite the existence of the problem considered. Interestingly, following a preliminary investigation, it remained unclear whether or not this is the case. To elucidate, the structural power possessed by the Gauteng Department of Education seems to be of significant utility to the resolution of the problem considered, but at the same time its lack of status power[16] suggests an inability to contain the ripples of effect caused by the problem considered [306]. In such cases of uncertainty, it is advisable to consider the aforementioned assumption invalid. As a result, if this were an actual analysis, it would be necessary to return to Phase 3 of the methodology in order to reconsider the affected.

In conclusion to the discussion on the execution of Phase 4, note that although the questions in the power analysis checklist are designed to elicit either a "yes" or a "no" in response,

[16] In this case, status power represents financial means.

the power analysis should not be viewed as some form of boolean analysis in which the only element of interest is the final output. The actual decision surrounding an entity's inclusion or exclusion should be based upon the relative measure of that entity's power possessed (and effect experienced) with respect to that of all other potential stakeholders, and not on a number of conceptual criteria of relevance alone. The power analysis checklist is simply meant to ground the analysis of an entity's power and act as an initial filter separating potentially relevant from irrelevant entities based on the measure and form of power that they possess.

All that remains unaddressed with respect to the hypothetical application of the methodology proposed to facilitate the identification of stakeholders is the execution of Phase 5. There is, however, not much to discuss in this respect. Essentially, it encapsulates the process of determining which entities are included as stakeholders to the pending endeavour based on the analyses conducted during both Phases 3 and 4 of the proposed methodology — a process that is predominantly at the mercy of an analyst's subjective interpretation of the information elicited during the aforementioned analyses. The only technicality worth reiterating before closing this section is the following: Entities severely affected by the problem considered are deemed relevant regardless of the measure of power they possess.

10.3 The state of a system

The evaluation of the methodology proposed to facilitate the identification of stakeholders, briefly summarised in the section prior, did more than simply reveal some of its inherent limitations — it brought to light a potentially detrimental phenomenon not yet addressed in the proposed[17] meta-methodology. That is, the *state of the system* containing the problematic situation considered. In hindsight it is clear that in its current form, the meta-methodology proposed is underlain by the assumption that real-world systems are characterised by what is here called a *state of order*, but as evident from the theoretical case study in respect of the fees must fall movement, real-world systems can also exist in a *state of chaos*[18]. These state classes are subsequently further elaborated upon, but in essence, a state of order implies that the entities in a system are able to coexist in the spirit of *collaboration*, while a state of chaos implies that the entities in a system are unable to do so. In order to address the manner in which the state of a system complicates the resolution of problems that reside therein, the meta-methodology is expanded to include a methodological phase called *Moderate system state*, as depicted in Figure 10.2. In the remainder of this section, this phase is elaborated upon.

Arguably, the most appropriate question with which to open the discussion on the newly added methodological phase, is why the phase in question is, in fact, beneficial. The answer to this question contains both a methodological and a practical component. The methodological reasoning behind the inclusion of the additional phase essentially comes down to the fact that the final phase of problem formulation (*i.e.* problem structuring) is markedly facilitated when conducted in the context of a system that exists in a state of order. This is made evident by

[17] The phenomenon in question is seemingly also not addressed in existing system solution methodologies in general.

[18] There is a third state in which real-world systems may exist, called a *state of disintegration*, which is excluded from the above typology as the meta-methodology proposed will never be employed in the context of such a system. A state of disintegration represents the case in which a problematic situation has been left unattended for such an extended period of time that the system containing it has disintegrated, thus terminating its existence. If the original system no longer exists, the problem contained within it is no longer addressable, and so a system in such a state is excluded from the discussions that follow. As a distinguishing criterion, note that it is only possible for a system to enter a state of disintegration if there exists a higher-level system into which such a system can disintegrate.

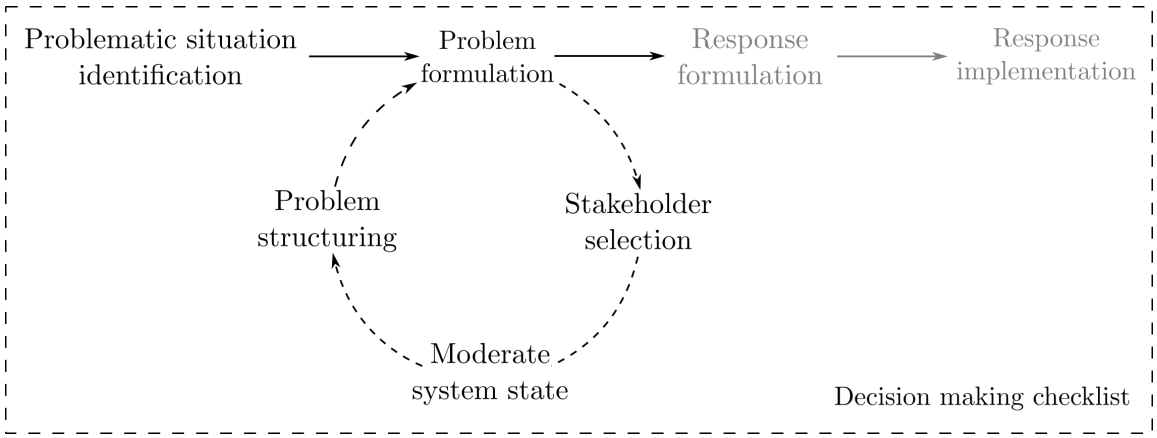


FIGURE 10.2: *A high-level overview of the phases within the proposed methodology with an updated lower-level depiction of the phases that constitute the process of problem formulation.*

formally defining the notion that governs the distinction between order and chaos — the notion of *collaboration*.

There are several definitions of the term collaboration in the academic literature: Westley *et al.* [480] defined collaboration as the process through which relevant stakeholders, who each perceive different aspects of a problem, can constructively explore their differences and search for solutions that go beyond their own limited vision of what is possible. Roberts *et al.* [384] similarly defined collaboration as an interactive process characterised by a trans-mutational purpose, voluntary membership, joint decision making, agreed-upon rules of interaction, and a temporary structure. Finally, Wood and Gray [489] defined collaboration as the process by which autonomous stakeholders engage in an interactive process (according to shared norms, rules and structure) in order to act or decide on issues related to the relevant problem domain. Although each of these definitions is valuable, none is entirely satisfactory within the context of this dissertation. Consequently, the following definition (building upon the above definitions) is proposed and employed: Collaboration exists among a group of autonomous stakeholders, relevant to the resolution of some problematic situation, if they are able to engage in the voluntary, interactive process of exploring the differences in their subjective understanding of the related problematic situation by employing shared rules, norms, and structures to facilitate the process of searching for solutions that go beyond the limited vision of any single stakeholder in a manner that is constructive and remains within the domain of civil debate.

In view of the above definition, the utility of conducting the task of problem structuring in a state of order is evident. In fact, the task of problem structuring, which involves gathering selected entities to investigate, debate and select a desired course of action, essentially presupposes such a state of affairs. That is, the methods of problem structuring (discussed in §9.2.2) are built upon the assumed utility of the social interaction of relevant entities, which is only possible if such entities are able to gather in the spirit of collaboration (*i.e.* if the related system is characterised by a state of order).

The practical motivation behind the methodological phase titled *Moderate the system*, on the other hand, is made salient by describing the mechanisms that underlie the presence of chaos in real-world systems. The first step in doing so, is illustrating that the presence of chaos, or its absence, is contained within the mechanisms of power, and furthermore, that there is a cost associated with a system in chaos (affecting all entities therein). During the discussion on the mechanisms of power in §9.2.1, the notion of a power-play was introduced and put forth as the medium by which entities in real-world systems are able to compete. This notion is once

again central as it underlies the manner in which systems transition between states of order and chaos. More specifically, the presence of chaos in a system is considered the product of what is here called *chaos-inducing power-plays* (*i.e.* power-plays that involve the imposition of a state of chaos as a fundamental mechanism in the manner in which they induce their influence). In particular, the coercive employment of either necessity or incapacitation as sources of power is uniquely associated with the induction of chaos. Consider the employment of necessity in its coercive form. In essence, this entails the removal of one's participation from a system in order to exhibit the extent of one's necessity to the proper functioning thereof, which is, by definition, chaos inducing. The employment of incapacitation in its coercive form, on the other hand, entails not only the removal of one's participation from a system, but the actual, intentional disruption thereof, which is, once again, chaos-inducing[19]. This description of the coercive employment of necessity and incapacitation illustrates why the induction of a state of chaos comes at a cost — chaos-inducing power-plays are fundamentally designed to inhibit the proper functioning of a system. The cost of chaos thus includes the loss of any benefit that would have been derived from the working of a system if it had remained in a state of order, as well as the cost of any damages caused to the system in question. The reality of such costs brings the discussion to the practical motivation behind the aforementioned methodological phase: Aiding analysts in minimising the cost associated with systems in chaos.

10.3.1 A technical enquiry

Before introducing the method proposed to facilitate the aforementioned task, it is beneficial to address some of the technicalities that underlie its design. In particular, the purpose of the subsequent discussion is to illustrate how non-chaos-inducing entities tend to aggravate the cost of a system in chaos by electing to adopt inappropriate strategies of competition in response to the employment of chaos by a competitor, and how the proposed method may facilitate counteracting this effect. To that end, a dynamic related to the employment of chaos-inducing power-plays is briefly explored below.

In the section prior, it was explained that a system in chaos incurs a certain cost, and furthermore, that this cost is not limited to non-chaos-inducing entities — taking a system into, and keeping a system in, a state of chaos also comes at a certain cost. The greater this cost, the more difficult it is to do so[20]. The attentive reader may thus wonder: Why would entities choose to employ chaos-inducing power-plays, or for that matter, any power-play whose employment comes at a certain cost? An entity will only follow this course of action if it deems the cost of doing so worth the potential benefit. In other words, if an entity in some real-world system surmises that the benefit[21] attainable from the system in question is utterly inferior to that merited by its power, and elects to rectify this perceived imbalance. It will only draw upon more *radical* power-plays (*i.e.* power-plays that employ chaos as a means) if the potential benefit of doing so is considered to outweigh the cost. The decision of whether or not to employ a certain power-play (*i.e.* whether or not to take a system into a state of chaos and keep it there) thus constitutes a subjective cost *versus* benefit analysis[22].

[19] Note that a power-play is only considered chaos-inducing if such a state of affairs is fundamentally linked to the manner in which it exerts its influence — the chaotic aftermath of an unsuccessfully employed power-play does not make it chaos-inducing. Accordingly, all other power-plays can be executed in the domain of order.

[20] This cost includes the loss of any benefit a chaos-inducing entity would have experienced if it had allowed the related system to remain in a state of order, as well as any losses it had to endure to maintain such a state.

[21] Benefit represents the positive component of the notion of effect, defined in §9.2.1.

[22] Note that non-chaos-inducing power-plays do not necessarily become redundant in a system that has entered a state of chaos. It remains possible for entities to attempt to influence the system in which they are contained by adopting any power-play available to them, no matter the state of that system.

The aforementioned cost *versus* benefit analysis is of central importance to the method subsequently proposed as it is relevant not only to an entity attempting to decide whether or not to cause chaos in a system, but also to a non-chaos-inducing entity attempting to decide how to respond to a chaos-inducing competitor. More specifically, the method proposed in the section that follows is essentially designed to facilitate such a case-specific cost *versus* benefit analysis in order to encourage entities to select competitive courses of action that best suit their pursued outcomes. To illustrate the fact that entities do at times fail in this task, consider the subsequent discussion on the manner in which the governing bodies of the South African education system responded to the chaos-inducing protagonists of the fees must fall movement (introduced in §10.1).

As stated, the South African system of higher education existed in a state of chaos during the fees must fall movement. In other words, there existed, at the time, a chaos-inducing entity therein. Although the identity of this entity has yet to be described formally, in accordance with the review of the fees must fall movement presented in §10.1, the entire proportion of students in the South African education system that partook in the fees must fall movement is, for the purpose of this discussion, considered the chaos-inducing entity therein[23]. To illustrate that the governing bodies of the South African education system failed in their response to the aforementioned entity, consider the excerpt below from a report published by the Centre for the Study of Violence and Reconciliation dedicated to the documentation of how select universities[24] responded to the chaos-inducing protagonists of the fees must fall movement [277].

> "In all nine chapters in this report, key informants asserted that university management was highly unresponsive to their demands. This involved Vice-Chancellors refusing to engage with student leaders or not coming to scheduled meetings. In her chapter, Sandile Ndelu narrates how the failure of the Vice-Chancellor at CPUT to arrive at a meeting made students angry, resulting in them rioting and burning university property. Similar stories were shared about TUT management being aloof and distant in their engagement with student leaders. However, some Vice-Chancellors, such as those at Rhodes and UCT, tried their level best to actively engage with

[23] Toward illustrating that their use of chaos is in harmony with the above discussion dedicated to why entities may elect to utilise chaos-inducing power-plays, consider the subsequent description of the power dynamics at play within the system of higher education in South Africa prior to the inception of the fees must fall movement.

Like many countries around the world, the South African system of higher education is structured so that the students contained therein typically need not partake in the power battle that exists therein. To elucidate, in accordance with the constitution adopted in 1996 [119], certain entities within the South African government are officially granted the formal purpose and power to act as representatives of the country's student body in order to ensure that they are treated in accordance with the statutes of that constitution. Importantly, although this strategy has the potential to truly benefit a country's student body, it also facilitates their exploitation in the case where chosen representatives elect to abuse their positions for the benefit of their own informally pursued outcomes — the system in question leaves little room for students to influence the system that contains them *directly* (*i.e.* without having to rely upon the formal structures designed to represent them). In fact, such influence resides almost exclusively in the informal power sources of necessity and incapacitation, defined in §9.2.1. In the case of South Africa, the ability of the aforementioned representatives to exploit their position seems to have been an all too enticing temptation [84, 414]. As a result, more than twenty years after the adoption of the South African constitution, many of the promises made therein remained largely unrealised [332, 349, 350]. In view of such a state of affairs, it is not surprising that the students of the South African education system lost faith in the formal structures designed to represent them (as is evident from the fact that the fees must fall movement did, in fact, take place), and concluded that the potential benefit of employing the informal sources of power available to them, and that in a coercive form (thereby inducing chaos), outweighed the cost of utilising chaos as a source of influence.

[24] The following universities were considered: The Cape Peninsula University of Technology, the University of the Western Cape, Tshwane University of Technology, the University of Cape Town, the University of KwaZulu-Natal, Rhodes University, the University of Limpopo, the University of Zululand, and the University of the Witwatersrand [277].

> student leaders. However, divisions and splinter groups within the #FeesMustFall movement negatively affected some of these engagements. It appears that a common response by the university management in all the case studies was to get court interdicts against the protesting students. The law was therefore used to silence voices of dissent as many students were arrested for contravening court interdicts."

In view of this excerpt, it is evident that the governing bodies of South African universities failed to respond appropriately to the protesting students of the fees must fall movement. More specifically, they failed to respond in a manner that dampened the cost of chaos and increased the measure of benefit attained. Several researchers have argued this point [277, 290].

It is toward mitigating such instances of poor decision making that the methodological phase titled *moderate the system* was designed.

10.3.2 Moderate the system

Addressing the state in which a system exists is a notion quite unique in systems theory. The sections prior, dedicated to this notion, included, among other subject matter, a delineation of the criteria that underlie chaos, as well as an account of the mechanisms by which, and the reasons for which, an entity may elect to induce chaos. The purpose of the current section is to now build upon this body of knowledge and present a method capable of facilitating the task of developing an appropriate strategy of competition in response to a chaos-inducing competitor, thereby attempting to minimise the cost of chaos. In order to facilitate the presentation of this method, a typology of the strategies of competition considered to be available to an entity when selecting how to respond to a chaos-inducing competitor is presented:

(1) An entity can acknowledge the power possessed by a chaos-inducing entity and grant it the influence (benefit) its power merits (perhaps by suggesting a compromise between its own, and the outcomes pursued by the chaos-inducing entity).

(2) An entity can respond competitively by employing either selected power-plays or allowing the passage of time[25] so as to increase the cost of keeping the related system in a state of chaos. The purpose of such an action would be to coerce the chaos-inducing entity into either abandoning the pursuit of its outcomes by means of chaos, or simply settling for a lesser degree of benefit than originally demanded.

The method proposed contains two parts: Part 1 is dedicated to an enquiry into the power-dynamics that characterise a chaos-inducing entity, and Part 2, to the development of a strategy of competition that may be adopted in response to the entity in question. These parts are thus designed to be executed sequentially. Accordingly, consider the steps that constitute Part 1 of the proposed method below:

Part 1

1.1 Identify and briefly describe the chaos-inducing entity.

1.2 Establish a channel of communication with the identified entity.

[25] There is a cost associated with keeping a system in chaos. The longer an entity has to maintain such a state, the larger the associated cost — an entity can rarely afford to keep a system in a state of chaos indefinitely.

1.3 Ascertain its demands, and its willingness to participate in a structured resolution of the problem considered, given the establishment of a temporary agreement as interim solution[26].

1.4 Delineate the distribution of power in the system in question. Do this separately for each of the sources of power[27] defined in §9.2.1.

1.5 Evaluate the demands of the chaos-inducing entity in accordance with the power it possesses.

1.6 Describe the current cost of chaos to all entities relevant to the problem at hand.

1.7 Assess the chaos-inducing entity's ability to maintain a state of chaos in the system in question[28].

As stated, Part 1 of the proposed method, presented above, is designed to facilitate an in-depth analysis of the power-dynamics that characterise a chaos-inducing entity and should act as input to the execution of Part 2. Accordingly, consider the steps that constitute Part 2 of the proposed method:

Part 2

2.1 Gather the non-chaos-inducing entities that were selected as stakeholders to the problem considered (during the previous methodological phase) that also possess power relevant to the selection of how the chaos-inducing entity should be responded to.

2.2 In accordance with the information uncovered during the execution of Part 1, facilitate the selection between pursuing Strategy (1) or Strategy (2), presented above.

2.3 If Strategy (1) is selected:

2.3.1 Proceed to formalise a preliminary agreement to be presented to the chaos-inducing entity as soon as possible. This agreement should include, but should not be limited to: The extent to which its demands can be met, the period over which the preliminary agreement will remain valid, and the steps that are to be taken during that period in order to resolve indefinitely the problem considered.

2.3.2 Make contact with the chaos-inducing entity and present the preliminary agreement. This will likely be followed by a period of negotiation during which entities may debate the contents of the agreement.

2.4 If Strategy (2) is selected,

2.4.1 Specify the goal of responding competitively (*i.e.* the benefit to be attained).

[26] The purpose of the above method is not to resolve the problem considered, but simply to facilitate the establishment of a state of order in the relevant system as soon as possible. Accordingly, the purpose of the second part of Step 1.3 above, is simply to determine whether or not it is possible to establish a temporary agreement between the non-chaos-inducing entities and the chaos-inducing entities that will hold for a predetermined period, during which the next methodological phase of problem structuring can be executed in the context of order. The execution of this phase may then potentially facilitate the indefinite resolution of the problem considered.

[27] During the execution of Step 1.4, it is important that analysts aid entities in conducting as fair a review as possible. Power possessed in the form of necessity and incapacitation are easily undervalued.

[28] When executing Step 1.7, it is important to consider the historical development of the problem considered.

2.4.2 List and describe, in some detail, the power-plays that are available toward achieving these goals.

2.4.3 Proceed to describe the potential cost associated with each of the available power-plays. This should be done separately in respect of the effect these power-plays will have upon chaos-inducing and non-chaos-inducing entities.

2.4.4 Given the ability of the chaos-inducing-entity to maintain a state of chaos (assessed in Part 1), and the perceived manner in which given power-plays will effect the entity in question, review the likelihood of attaining the goals described in Step 2.3.1.

2.4.5 In accordance with the information uncovered during the steps prior, re-evaluate whether or not pursuing Strategy 2 is indeed appropriate to the situation at hand. If so, proceed to select from the available power-plays those that are best suited to the manner in which the gathered entities wish to compete[29].

2.4.6 Proceed to formalise a detailed description of the steps involved in the execution of the selected power-plays as well as a time-line of their execution.

2.4.7 Finally, specify a predetermined point in time at which the effectiveness of the selected power-plays will be reviewed with the goal of determining whether or not their employment was, in fact, worthwhile.

In view of the above method, it is acknowledged that there is a slight overlap between the two strategies of competition utilised. To elucidate, it is possible that during the execution of Strategy (1), the task of establishing a preliminary agreement satisfying both non-chaos-inducing and chaos-inducing entities is found to be infeasible as neither set of entities are willing to compromise to the extent desired by their competitor. In such a situation, non-chaos-inducing entities may elect to pursue Strategy (2), albeit purely for the purpose of coercing a chaos-inducing competitor into settling for a lesser degree of benefit.

In conclusion, although it is certainly important to establish a state of order in the system under consideration as soon as possible, practitioners are cautioned against attempting to do so prematurely. In order to address the presence of chaos in a system, for example, its source, in the form of chaos-inducing entities, must first be identified, and then the collaboration of those entities must be achieved, which typically involves the establishment of some form of temporary agreement. Consider the wisdom of establishing such an accord without first having identified all entities relevant to its establishment. In such a case, it is highly likely that the accord may be deemed unacceptable at a later point in time by several stakeholders not initially considered — resulting in a conundrum. As such, although it is important to address the presence of chaos in a system as quickly as possible, this phase of the proposed methodology should be conducted with care.

10.4 Chapter summary

This chapter was dedicated to an identification of the shortcomings in one of the methods presented in the meta-methodology of Chapter 9 during its hypothetical application to a historical, real-world endeavour. To that end, the context of evaluation was briefly delineated in §10.1, and this was followed in §10.2 by an account of the hypothetical application of, and the improvements made to, the method in question. Significantly, the application of the aforementioned

[29] If more than one power-play is selected, it is necessary to ensure that their simultaneous execution is, in fact, feasible.

method facilitated the identification of a problematic phenomenon not addressed in the initial version of the meta-methodology proposed. Accordingly, an additional methodological phase titled *moderate the system*, designed to address this phenomenon, was presented in §10.3.

CHAPTER 11

An expert evaluation

Contents

In accordance with the research methodology adopted in this dissertation, the development of the proposed meta-methodology necessitated a second stage of evaluation — the presentation thereof to an expert analyst. The aim of this stage of evaluation is to elicit potential shortcomings in the methodology that remain despite its hypothetical application in the context of a theoretical case study, as presented in Chapter 10.

The chapter opens in §11.1 with a brief account of feedback received from an expert analyst in respect of the working of the proposed meta-methodology, and this is followed by the presentation of an updated version of the methodology in accordance with this feedback, in §11.2. The chapter then comes to a close with a short summary in §11.3.

11.1 Expert feedback

The lion's share of the critique received from the expert analyst (formally introduced in Chapter 2) concerns the design format of the methods proposed to facilitate each of the three constituent phases of the task of problem formulation. In general, the meta-methodology was considered to be of noteworthy utility, but the expert analyst warned that unless a schematic representation of its constituent parts is developed, the value thereof could be lost on future practitioners[1]. That is, in its current form, the value of the methodology is somewhat overshadowed by the detailed fuzziness of its presentation.

The expert analyst provided two additional points of concern, one related to the method proposed to facilitate the selection of stakeholders, originally introduced in §9.2.1, and the other related to the decision-making checklist, originally introduced in §9.3. In respect of the first of these methods, the expert analyst adjudged the fact that no guidelines are provided to facilitate the practical delineation of a functional network to be a shortcoming that should be addressed, and in respect of the second, the expert analyst considered one of the sub-methods

[1] The decision-making checklist, introduced in §9.2.1 was exempt from this criticism.

contained therein to be somewhat out of place. To elucidate, the final component of the decision-making checklist contains the suggestion that if the polarising nature of group decision making is suspected to underlie a difficulty experienced in the establishment of consensus among competing entities, it may be beneficial to request that all entities select a single representative of their cause, and then organise an "executive meeting" among the selected individuals (see §9.3 for more information). The analyst brought to light an inconsistency between the formal description of the decision-making checklist as a framework designed for consideration at strategic junctures during an endeavour, and the case-specific nature of the aforementioned method contained therein. Accordingly, it was suggested that this final method is removed.

Before proceeding to a discussion on the improvements made to the proposed meta-methodology, a high-level overview of its constituent phases is again presented here toward facilitating the reader's comprehension of the components therein discussed subsequently, as depicted in Figure 11.1

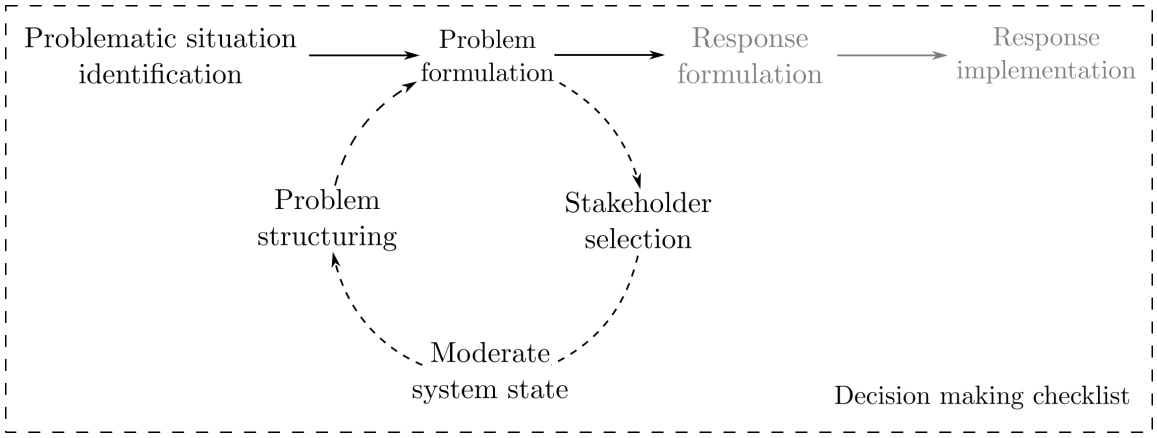


FIGURE 11.1: *A high-level overview of the phases within the proposed meta-methodology.*

11.2 Improvements made

In response to the respective points of concern identified by the expert analyst, several modification are made to the proposed meta-methodology. Arguably, the most substantial modification is the formulation of a schematic representation of each of the constituent methods of the task of problem formulation. The first such method, dedicated to the task of stakeholder selection, is depicted in Figure 11.2. As can be seen in the figure, the following colour scheme is utilised to distinguish different forms of information: *Crayon-pink* designates the methodological phase considered, *burlywood-brown* highlights its constituent parts, and *peachpuff-orange* indicates supplementary information. Notably, one of the components presented in the figure, called *Constructing a functional network*, is included only in response to the second point of concern of the expert analyst. That is, the lack of guidance given in respect of how a practitioner may go about visually organising the development of a functional network. In order to rectify this shortcoming, an exemplar of how this task may be accomplished is presented (a more detailed demonstration of the technique is provided later in this dissertation).

In addition to the aforementioned exemplar, the components of the method depicted in Figure 11.2 detail the steps involved in its execution, along with several pieces of supplementary information. The first component, depicted in the top left-hand segment of the figure, consists of five steps that underlie the method proposed. In the order presented, these steps involve

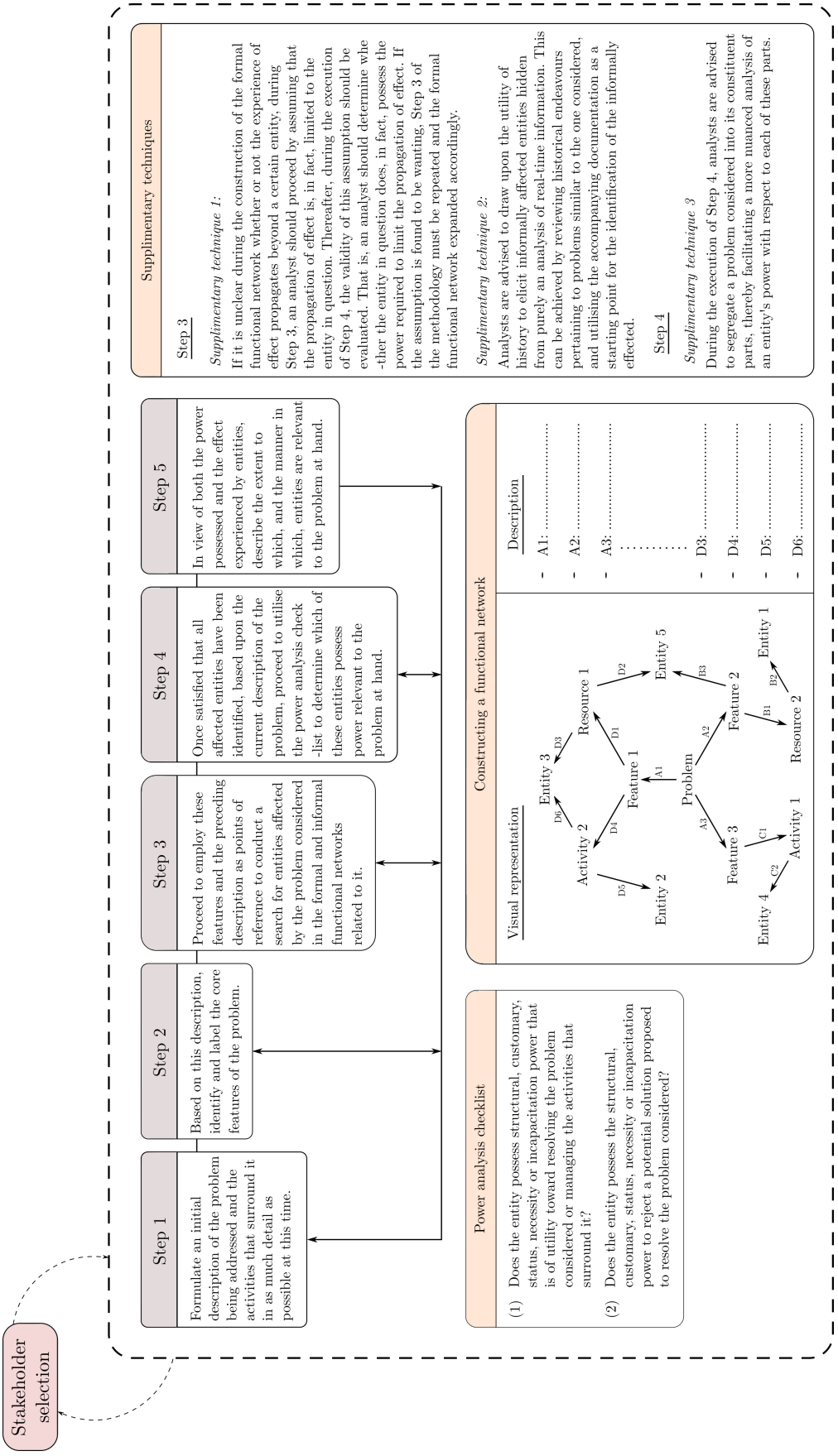


FIGURE 11.2: *A schematic representation of the method titled Stakeholder selection.*

developing an initial description of the problem considered, delineating the features that characterise it, identifying the entities affected thereby, distinguishing the subset of such entities that actually possess power relevant to its solution, and describing the nature of this relevance, respectively[2]. The second component, depicted in the bottom left-hand corner of Figure 11.2, is dedicated to a summary of the properties that qualify the relevance of an entity's power (utilised in the execution of Step 4), and the third, depicted along the right-hand border of Figure 11.2, contains a description of three supplementary techniques that facilitate the execution of the method proposed.

As the second phase in the formulation of real-world problems, titled *Moderate the system*, is included in the scope of the expert analyst's first point of critique, a visual representation of its working is presented in Figure 11.3. As can be seen in the figure, the method contains two parts, one dedicated to an investigation into the nature of a chaos-inducing entity, and the other, the design of an appropriate response to such an entity. The first part is made up of seven steps which, in the order depicted, involve identifying the chaos-inducing entity, establishing a channel of communication with the entity in question, formulating an understanding of its demands, investigating the power dynamics that characterise the related system, evaluating the extent to which the demands of the entity in question are proportional to its power, describing the cost of chaos in respect of all the entities affected thereby, and assessing the chaos-inducing entity's ability to maintain such a state, respectively.

As stated in §10.3.2, the steps involved in the second part of the method in question depend upon the general class of competitive strategy selected by the non-chaos-inducing entities. Notably, the two strategy classes that are considered available to such entities are depicted along the right-hand border of Figure 11.3 in the form of supplementary information (*i.e.* Strategy 1 and Strategy 2, as described in §10.3.2). If Strategy 2 is selected, part two of the method proposed involves specifying the goal of responding competitively, describing the power-plays that are considered of utility to the stated goal, assessing the cost associated with their use, re-evaluating the appropriateness of pursuing Strategy 2, selecting the power-plays that are to be utilised, formalising a detailed description of their execution, and finally, specifying a predetermined point in time at which the effectiveness of the selected power-plays are to be reviewed with the purpose of determining whether their continued use is, in fact, worthwhile. If Strategy 1 is selected, on the other hand, part two of the method reduces to formalising a preliminary agreement that is to be presented to the chaos-inducing entity specifying (among other details) the extent to which its demands can be met, the period for which the agreement is valid, and the steps that are to be taken during this period in order to resolve indefinitely the problem considered.

The work presented in respect of the methodological phase titled *Problem structuring*, is unique in the context of the proposed meta-methodology. The contribution made therein does not take the form of a method proposed, but rather, a review of a collection of problem structuring methods sourced from the academic literature. The purpose of reviewing the methods in question is to inform practitioners of the sociological paradigm upon which their designs are based and thereby facilitate their case-specific selection. Accordingly, a schematic representation of the work presented in this regard is portrayed in Figure 11.4. As may be seen in the figure, this includes a description of the sociological paradigms of interest, the respective methods related thereto, and a discussion of the weakness associated with each paradigm. More specifically, a review of the *interpretive* and the *radical humanist* paradigms, along with the methods contained therein, are presented in the left-hand and right-hand segments of Figure 11.4, respectively, while

[2]These activities are to be executed in sequence, but may be returned to and revised *ex post*.

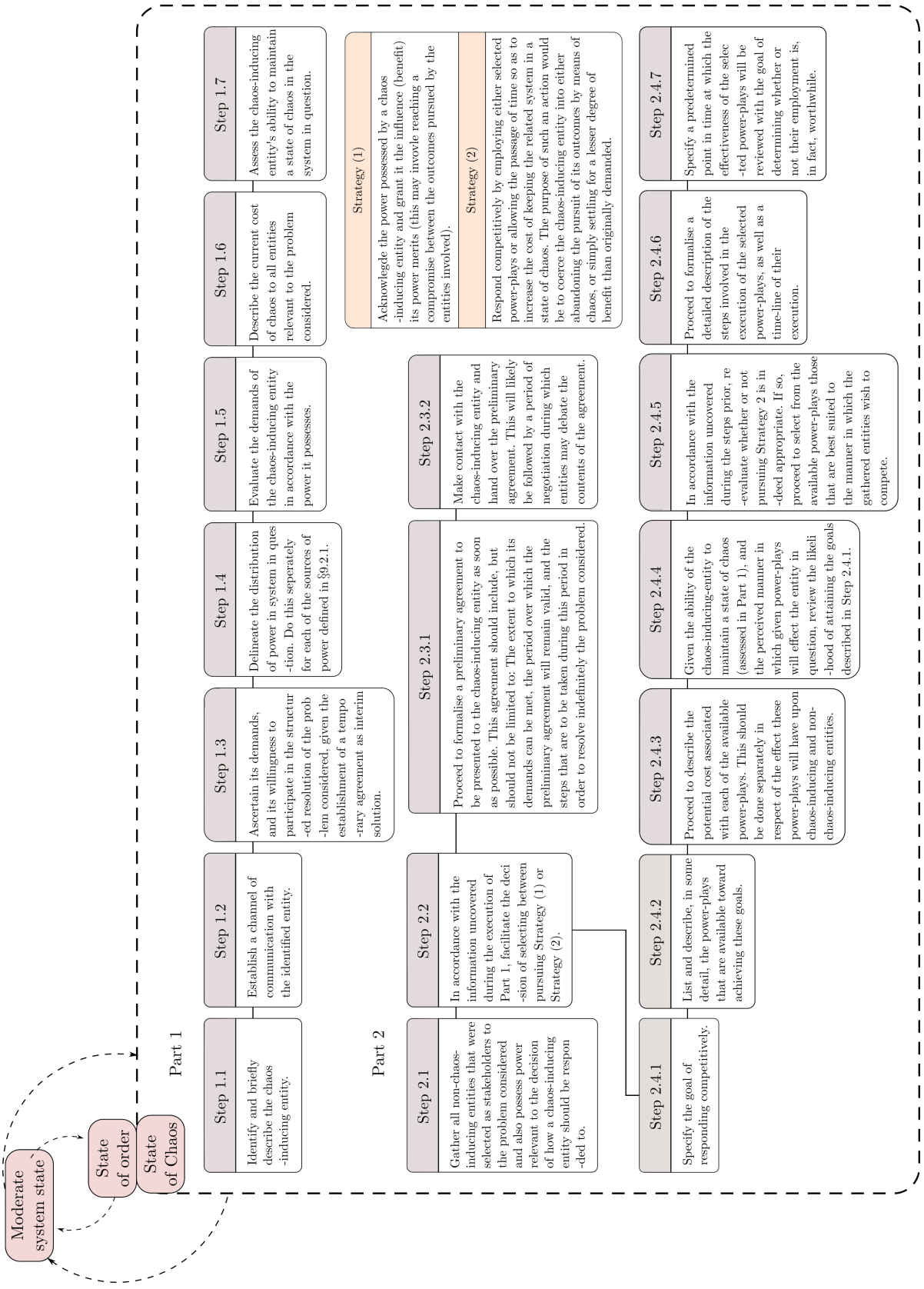


FIGURE 11.3: *A schematic representation of the method titled Moderate the system.*

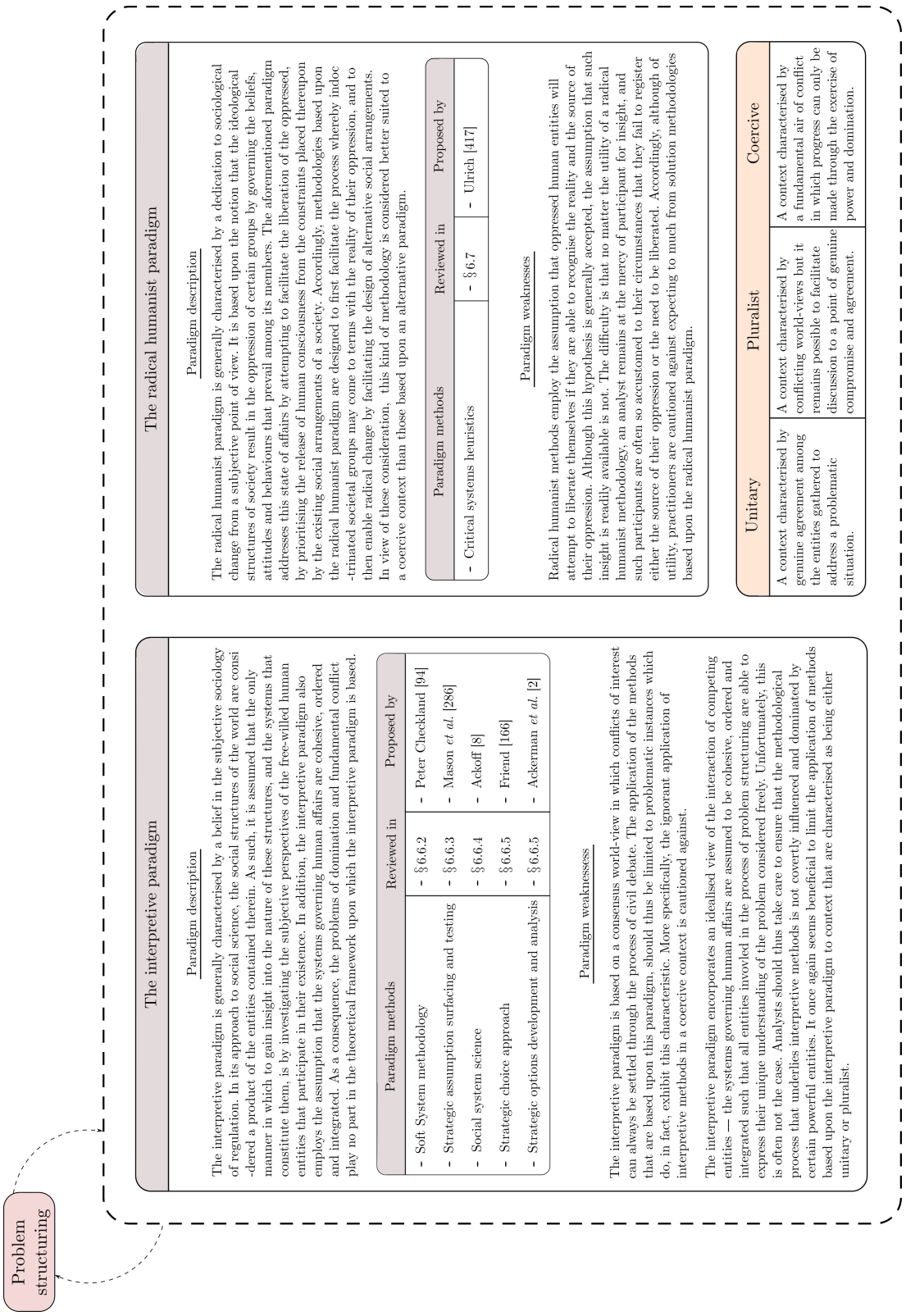


FIGURE 11.4: *A review of a collection of problem structuring methods.*

TABLE 11.1: *A decision-making checklist.*

Dynamics addressed	Architecture
Anchoring bias & Affective failures	The tendency to satisfice, even when significantly better solutions may be attained with little further effort.
	Reconsider the selected decision options. In what way can these courses of action be further improved? Are these improvements worth pursuing?
Affective failures	The tendency to downplay the consequences of a decision option if those consequences are not personally experienced, or if they are realised over an extended period of time.
	Reconsider the selected decision options. What are the long-term consequences of the various decision options? Describe the worst-case scenario. Are these outcomes acceptable?
Optimistic bias, Illusion of control & Misconceptions of chance	The tendency to be overly optimistic about the future and one's ability to realise outcomes pursued.
	Reconsider the selected decision options. Make a list of the risks and uncertainties associated with the respective options. Are these risks appropriate? What steps can be taken to render the future less volatile?
Illusion of validity, Insensitivity to predictive accuracy & Confirmation bias	The tendency to prematurely accept the validity of sourced information if that information coincides with one's preconceived beliefs, while simultaneously neglecting information that contradicts those beliefs.
	Reconsider the information upon which the selected decision options are based. Is the validity of this information apparent? Formally state/restate the information that contradicts the appropriateness of the selected decision options. On what grounds is this information rejected?
Groupthink, Abilene paradox & Shared information bias	The tendency to withhold information believed to be contrary to the opinion held by the dominant players in a context, particularly when that information is possessed only by a single individual.
	Reconsider the various decision options. In an anonymous fashion, request that participants describe any issue related to the realisation of the various decision options that have not yet been addressed sufficiently.

a typology of dissimilar problem types is presented in the bottom right-hand corner of the figure in the form of supplementary information.

In order to rectify the noted inconsistency in the decision-making checklist, the expert analyst suggested that one of the methods contained therein be removed. That is, the method designed to facilitate the mitigation of the unwanted effect of group polarisation and the discontinuity effect. As such, an updated version of the decision-making checklist, in accordance with this suggestion, is presented in Table 11.1.

11.3 Chapter summary

In order to refine the meta-methodology proposed, it was subjected to the critique of an expert analyst. This second stage of evaluation was the topic of interest in this chapter. Accordingly, the chapter opened in §11.1 with a brief account of the points of concern identified by the expert analyst, and this was followed in §11.2 by a discussion on the resulting modifications made to the methodology proposed.

The work presented in this chapter should not be viewed in isolation. That is, the modified version of each of the constituent parts of the methodology presented remains linked to, and dependent upon, the more detailed descriptions of their working in chapters prior. Practitioners are thus cautioned against attempting to apply the version of the methodology documented here without first reviewing the work documented in chapters prior.

CHAPTER 12

A practical case study

Contents

The meta-methodology proposed in this dissertation is most accurately described as a guideline designed to support practitioners in the solution of real-world problems. As such, the employment thereof does not guarantee successful problem solving — its purpose is limited to a role of support. In a final attempt to evaluate the utility of the proposed meta-methodology in this capacity, this chapter is dedicated to a documentation of its practical application in the context of the South African energy sector.

The chapter opens in §12.1 with a brief introduction to the problem considered, and this is followed in §12.2 and §12.3 with an account of the practical application of the meta-methodology proposed. In §12.4, the chapter finally comes to a close with a brief summary.

12.1 Background to the problem considered

In the history of electricity production, the South African power utility, Eskom, was long held in high esteem [413]. Today, however, the pendulum of public opinion has swung to, and has remained in, the opposite extreme. The variability of Eskom's power production capabilities is now considered one of the most significant challenges facing the South African economy [340]. In the words of the South African Public Enterprises Minister Pravin Gordhan [396], "Its going to be a huge struggle for us to overcome this emergency."

Due to the state of electricity production in South Africa, its distribution systems have, since the turn of the 21st century, witnessed a rapid increase in what are called *small-scale-embedded-generation* (SSEG) facilities, consisting primarily of solar and wind energy generation (*e.g.* roof-mounted solar panels on houses) [237]. Formally, entities with SSEG capabilities are called *prosumers* as they are able to switch between consuming and producing electricity as required [294]. That is, in addition to supplementing their own energy needs, prosumers contribute to alleviating the general energy load of a region by feeding the energy they produce, but do not use, back into the *grid*[1], in exchange for a small *feed-in-tariff*[2], so that distribution network operators can

[1] The term *grid* is synonymous with electricity distribution network.

[2] In 2018, for example, the Municipality of Cape Town paid individuals with SSEG capabilities a feed-in-tariff

redistribute and resell that energy for a profit [92]. This facility, the local trade of electricity, is one of the central themes investigated in this chapter. More specifically, the case study presented here involves a Cape Town-based enterprise[3], called Switch, and its founders' vision of implementing a particular form of peer-to-peer electricity trading in South Africa [443].

Switch was established in response to the work of Murray [334], who in his master's thesis investigated the utility of several electricity trading pricing strategies in the South African context and determined that a peer-to-peer blockchain-based dynamic fee trading strategy, where tariffs are determined in accordance with real-time supply and demand information [399], had the potential to aid South Africa in the recovery of its faltering energy system. According to Murray [334], such a pricing strategy would incentivise both electricity consumption during times of high production, and production during times of high demand, thereby aiding network operators in the difficulties associated with *grid-balancing* and *peak-load shifting* [194]. In 2018, Switch was established to make Murray's recommendation a reality [443].

At the point in time at which the case study documented here commenced, Switch had been a year in development and its founders had succeeded in developing an initial version of the software package required to facilitate the implementation of the aforementioned trading strategy. There were still, however, several technicalities that prohibited its official launch into the market, including, but not limited to, procuring the capital required to do so, and selecting the most viable energy trading business model to accompany the selected trading strategy. After a preliminary investigation, it was determined that the meta-methodology documented in this dissertation had the capacity to aid in the resolution of the aforementioned difficulties. More specifically, the first and final methods in the meta-methodology were deemed applicable. The method for the selection of stakeholders was applied in order to identify the entities, activities, and resources considered relevant to the avenues of future development available to Switch (in order to determine their validity), and the decision making checklist was applied in order to ensure that any decisions made in this regard were, in fact, appropriate. An account of the application of these methods is presented in the remainder of this chapter.

All case-related information discussed subsequently was gathered *via* personal communication with Switch's founders during the four-month period between February and May 2019 [335].

12.2 Identifying relevant stakeholders

As stated, the purpose of applying the method designed to facilitate the selection of stakeholders was to aid Switch's founders in overcoming the difficulties that inhibit them from establishing Switch as an official player in the South African energy sector. Accordingly, in fulfilment of the first step in the method in question, the problem considered[4] is described as follows:

> Switch exists as an informal enterprise. Its founders have developed a blockchain-based software platform capable of facilitating the local, real-time trade of electricity,

of 84.95 cents (ZAR) per kWh of electricity they fed back into the grid, which they then resold for 185.32 cents (ZAR) per kWh [116].

[3]Since the conclusion of the case study documented here, Switch has moved its base of operations to Johannesburg, South Africa.

[4]It is necessary to qualify the application of the method in question to the problem described above. That is, it is necessary to illustrate that the system which contains it is characterised by the property that governs the class of system for which the meta-methodology proposed in this dissertation is designed (*i.e.* that the entities contained therein are forced to coexist despite being in competition with one another). It should, however, be rather obvious that this is indeed the case as the objectives pursued by Switch's founders are fundamentally linked to the other entities that exist in the South African energy system.

> but are unsure how to proceed in respect of its future development. The difficulties they face include access to capital, uncertainty in the particular energy trading business model that best suits their enterprise, and the state-dominated nature of the South African energy sector.

In view of the above preliminary description of the problem considered, three features were identified that characterise it (in fulfilment of Step 2), namely *strategic value offering*, *financial needs*, and *renewable mandate*, as depicted in Figure 12.1. In the order listed, these features are representative of the technicalities surrounding the selection of an appropriate energy trading business model, the financial requirements related to Switch's development, and the renewable energy mandate built into Switch's value-offering.

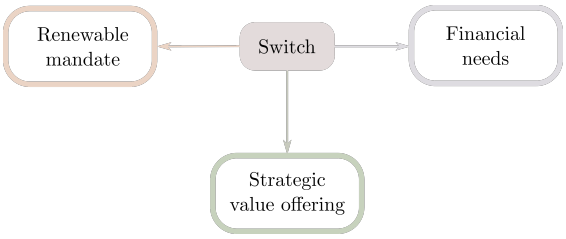


FIGURE 12.1: *A schematic depiction of the core features of the problem considered.*

The next methodological phase involved constructing the functional network induced by each of the features listed above. Notably, the network related to the problem considered is extensive and, as such, the level of detail in which feature-specific components are discussed is adjusted as deemed appropriate. With this in mind, the first feature of interest is the strategic value offering associated with the enterprise in question, or rather, the uncertainty that characterises it. As stated in §12.1, at the point in time when the case study documented here commenced, Switch's founders had already selected the energy trading strategy they wished to implement, but were struggling to settle on an energy trading business model that would actually facilitate its uptake in South Africa's energy market. Based upon a preliminary analysis, two energy trading models were deemed viable, namely *third-party trading* and *wheeling*. The essential difference between these models is that in the context of the former, Switch would become an energy trading entity (similar to foreign exchange traders) and actively participate in the trade of electricity (purchasing electricity from prosumers, transporting it through a distribution network and selling it to consumers) [334], while in the latter, Switch would simply function as a electricity trading software solution provider that facilitates the direct trade of electricity between consumers and prosumers within a particular distribution network [334]. In other words, both models are built upon the dynamic pricing strategy Switch incorporates, but in the one case, Switch earns its keep by purchasing and reselling electricity, while in the other Switch simply charges a monthly *wheeling fee* for the trading platform it provides. Although these differences may seem subtle, they result in distinct propagations of effect and, as a result, in an attempt to aid Switch's founders in an appropriate selection between trading models, the functional networks related thereto were constructed, as depicted[5] in Figure 12.2.

In accordance with the strategy of presentation adopted in this chapter, only the most important insights uncovered during the construction of the functional network are discussed. The reader is referred to Table A.1 in Appendix A for a detailed description of each of the labelled arrows in Figure 12.2. With this in mind, consider first the left-hand segment of the functional network in Figure 12.2, induced by the third-party energy trading model. As can be seen in the figure, the

[5]The coloured blocks in Figure 12.2 denote entities relevant to Steps 4 and 5 of the methodology applied.

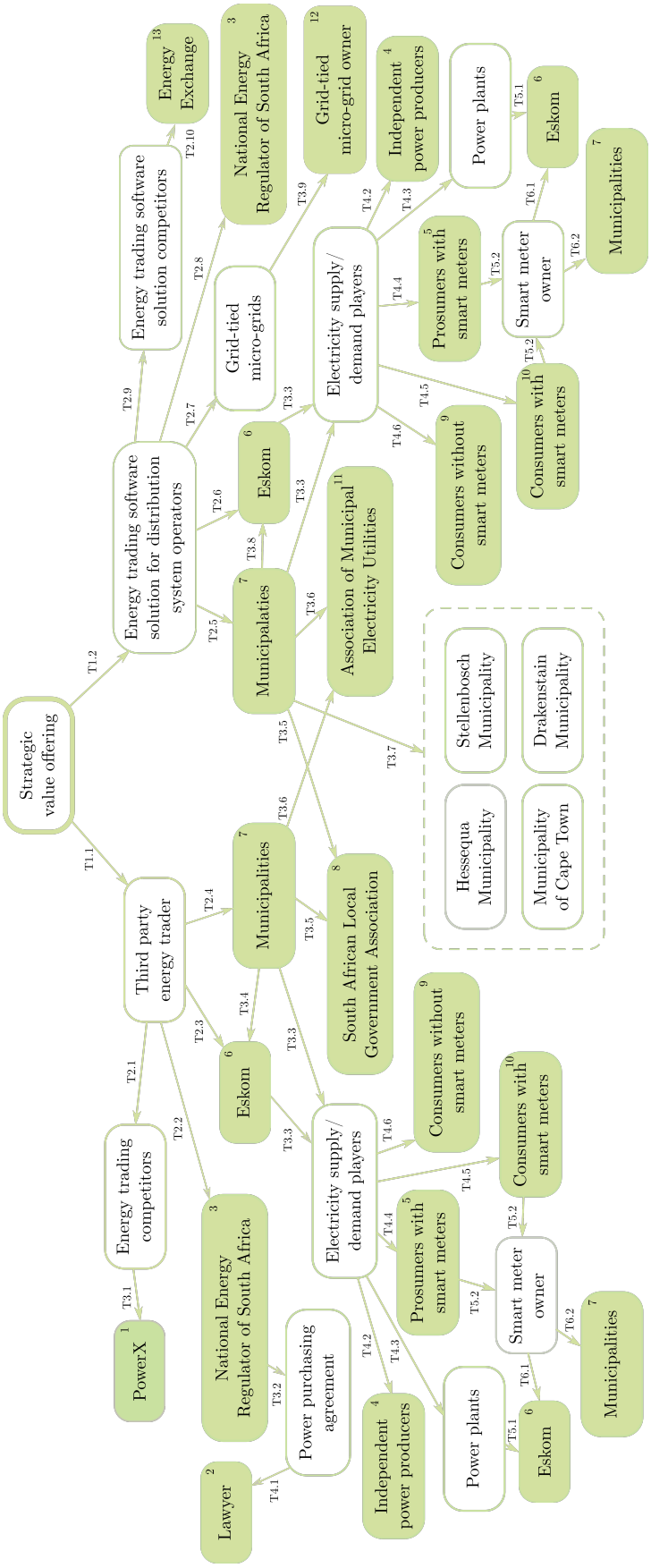


FIGURE 12.2: *A functional network induced by the strategic value offering associated with the problem considered.*

propagation of effect induced by this kind of business model occurs along four main paths (*i.e.* network edges). The first such path, labelled T2.1, corresponds with the reality that Switch will have to compete with, and thus affect, existing third-party energy traders[6], while the second, labelled T2.2, is representative of the legislative requirements associated with third-party energy trading in South Africa. More specifically, it was uncovered that in order to conduct business in the capacity of third-party energy trader, an enterprise must possess an appropriate trading license, in addition to getting the pricing strategy it utilises approved. The National Energy Regulator of South Africa [341] is the legal body governing these affairs and, as such, is considered formally affected, as depicted in the figure. The remaining edges considered primary, labelled T2.3 and T2.4, respectively, correspond to the utility of Switch's value offering as third-party energy trader in alleviating several of the difficulties that plague both Eskom and municipal distribution networks.

The right-hand segment of the functional network depicted in Figure 12.2, induced by the second energy trading model considered, is in many ways similar to the segment of the network described above. There are, however, several key differences. For one, although the National Energy Regulator of South Africa is again considered affected (see link T2.8), this time the propagation of effect concerns only the approval of the dynamic pricing strategy built into Switch's software package — no third-party energy trading licence is required. Additionally, the propagation of effect from the model in question now occurs along five main edges (not four). That is, an additional edge, labelled T2.7, which denotes the propagation of effect to what are called *grid-tied micro-grids* is now included. Formally, a grid-tied micro-grid is defined as a localised group of electricity producers and consumers that exist behind a single municipal meter and operate either connected to, or disconnected from, the general distribution network, as physical or economic conditions dictate (*e.g.* an estate with SSEG capabilities) [60]. In the third-party energy trading model, trading within this kind of micro-grid was deemed infeasible due to the economies of scale required to function as a third party energy trader [334]. In the wheeling trading model, however, this technicality is no longer relevant and, as such, grid-tied micro grids are included in the functional network related thereto. A final key difference between the network in question and the one described above, is that during a preliminary market analysis several municipalities were identified to be interested in the dynamic fee trading solution Switch will be able to provide if it elects to adopt the wheeling business model (see link T3.7).

As may be seen in Figure 12.3, the functional network induced by the second feature deemed core to the problem considered (*i.e.* the financial difficulties that characterise it) is slightly less involved than the one discussed prior. That is, seven entities were deemed affected on the basis of five potential ways of going about raising capital, namely acquiring a lone, procuring venture capital, organising an initial coin offering[7], joining a business incubator, or partnering with another organisation. Once again, the reader is referred to Table A.1 in Appendix A for a detailed description of the labelled arrows in Figure 12.3.

The final feature related to the problem considered (*i.e.* renewable energy mandate), is rather unique in the sense that instead of corresponding with a particular difficulty experienced, it simply characterises one of the outcomes pursued by Switch's founders — the sustainable development of the South African energy sector. As such, the functional network induced thereby contains several entities that operate in the South African context and share in this pursuit. As depicted in Figure 12.4, the entities deemed affected include several non-profit support organisations, as well as the South African Department of Energy (see Table A.1 in Appendix A for a detailed description of each of the labelled arrows in the figure).

[6] Interestingly, only one such enterprise could be identified, PowerX, a Johannesburg-based energy trader [367].
[7] An initial coin offering is a type of capital funding using cryptocurrencies [109].

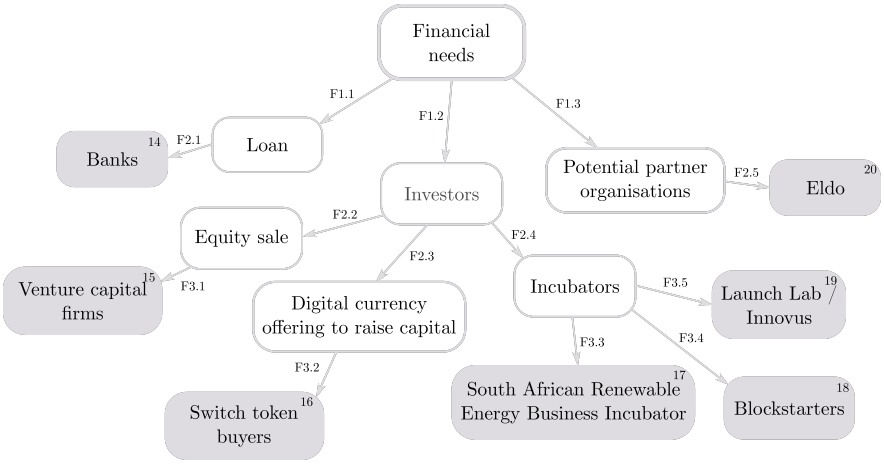


FIGURE 12.3: *A functional network induced by the financial requirements associated with the problem considered.*

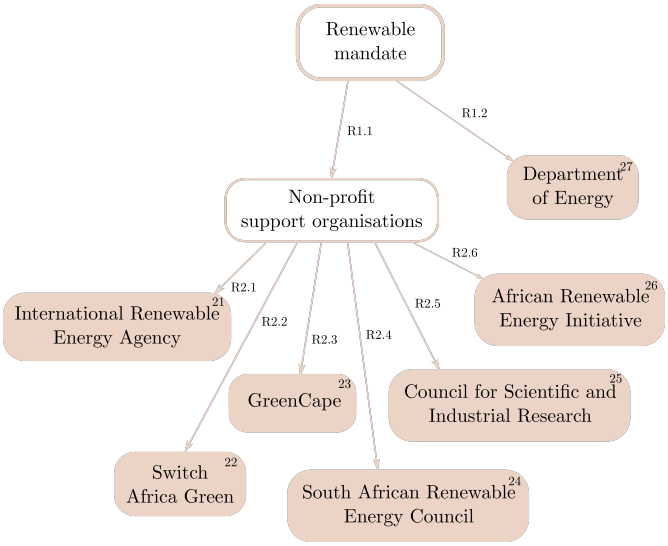


FIGURE 12.4: *A functional network induced by the renewable mandate associated with the problem considered.*

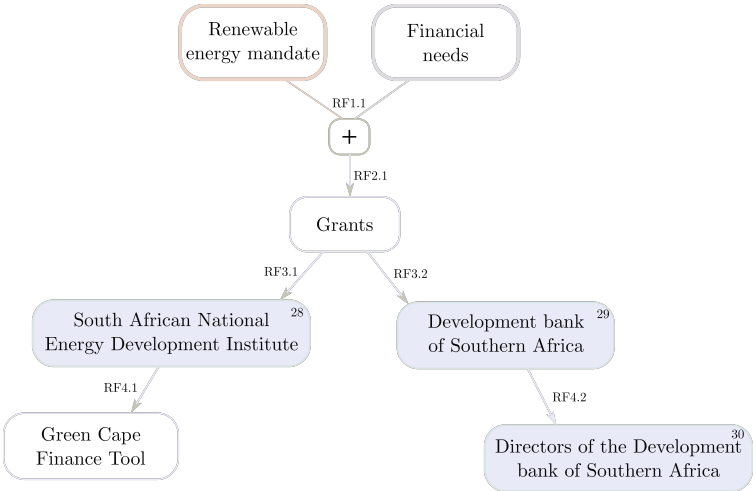


FIGURE 12.5: *A functional network induced by the combination of the financial requirements and the renewable mandate associated with the problem considered.*

Before continuing to the next phase of the analysis, there is a final segment in the functional network related to the problem considered that has to be discussed. During the execution of Step 3 it was determined that the combination of the financial and the renewable qualities described above results in a unique subset of affected entities. More specifically, several organisations were identified to invest only in technology-driven innovations specifically aimed at furthering South Africa's sustainable development. As may be seen in Figure 12.5, this includes the *South African National Energy Development Institute* and the *Development Bank of Southern Africa*[8]. The reader is again referred to Table A.1 in Appendix A for a detailed description of each of the labelled arrows in the figure.

The remainder of the analysis consisted of two parts, Step 4, aimed at investigating the nature and measure of the power entities deemed affected possess, and Step 5, aimed at determining whether, and if so, why the power these entities possess is, in fact, relevant. Although executed in sequence, the aforementioned methodological steps are presented in tandem. That is, for each entity deemed affected, an analysis of the extent, and the relevance of its power is presented in Table 12.1. As may be seen in the table, the first column contains a numeric label that corresponds with the numeric tag placed in the top right-hand corner of each of the coloured boxes in the functional network presented in Figures 12.2–12.5, while the second and third columns are dedicated to a description of the nature, and relevance of entities' power, respectively. Note that where applicable, the evaluation of entities' power is conducted separately for the two trading models considered.

Based upon the enquiry documented in Table 12.1, a total of fourteen entities were selected as primary stakeholders to the problem considered[9]. That is, the National Energy Regulator of South Africa, prosumers with SSEG capabilities, consumers equipped with smart meters, Eskom, local municipalities, grid-tied micro-grid owners, Eldo, Blockstarters, the South African Renewable Energy Business Incubator, the Council for Scientific and Industrial Research, PowerX, Energy Exchange, the Development Bank of Southern Africa, and finally, one of the directors of the Development Bank of Southern Africa.

In accordance with the power analysis checklist, originally presented in §9.2.1, the aforementioned entities were selected as stakeholders either because they possessed the power to inhibit the implementation of a solution proposed to the problem considered, or because they possessed power of considerable utility thereto. To elucidate, the National Energy Regulator of South Africa was selected as stakeholder as it is responsible for adjudicating both third-party energy trading licence applications and the validity of newly proposed energy trading pricing strategies. Prosumers and consumers were included as they are fundamentally linked to Switch's success — unless they are interested in Switch's value-offering, it will fail. Eskom and local municipalities, on the other hand, were selected because of the measure of control they possess over the South African electricity distribution network, which, interestingly, grants them the ability to either frustrate or facilitate the pursuit of Switch's objectives. Grid-tied micro-grid owners were deemed stakeholders as they are proprietors of a context seemingly well-suited to the pursuit of Switch's future development, while Blockstarters and the South African Renewable Energy Business Incubator were included due to their expertise in blockchain-based and renewable energy driven business development, respectively. Of the entities that remain, Eldo was selected as stakeholder because of its potential as partner-firm, the Development Bank of Southern Africa was selected because of its capacity to provide capital funding, one of its directors was included

[8]Significantly, one of the directors of the Development Bank of Southern Africa was identified as being a personal acquaintance of one of Switch's founders.

[9]This does not imply the that remaining entities are irrelevant, simply that a selection had to be made.

TABLE 12.1: *An analysis of the power possessed by, and the disposition of, the entities deemed affected during Step 3 of the methodology.*

label	Power description	Power Relevance
1	Status power in its operational infrastructure (*e.g.* developed hardware and software), clientele, brand reputation, and domain-specific knowledge.	This entity may be approached as a source of insight into the difficulties of operating as a third-party energy trader.
2	Customary power in its legislative authority, and status power in its domain-specific knowledge.	Legally, a lawyer has to certify the power purchasing agreement Switch's founders must submit in order to qualify for a third-party energy trading licence.
3	As the legal body responsible for adjudicating whether or not Switch may operate as a third-party energy trader and/or whether it may implement the dynamic fee trading strategy it proposes, this entity possesses considerable structural power.	**Third-party trading:** Unless this entity approves Switch's trading licence application and grants it permission to implement the fee structure it utilises, Switch will not be able to enter the energy market as a third-party energy trader. It is worth noting that there seems to be significant difficulty in attaining the licence in question. Of the companies that have applied, only one has been successful [334]. **Wheeling** Notably, if Switch pursues this option, no trading licence will be required, but this entity will still have the power to disallow Switch from implementing a dynamic-fee pricing strategy.
4	This entity possesses considerable necessity power as Switch's success depends upon its willingness to adopt the aforementioned dynamic fee pricing strategy.	Unless this entity is interested in the business model Switch's founders select, the likelihood of Switch's success decreases substantially.
5	Unless this entity elects to utilise the energy trading platform Switch's founder has developed, its vision of an energy trading South Africa will fail (no matter the trading model its founders select). Consequently, the entity in question possesses considerable necessity power.	This entity is considered a gatekeeper to the outcomes Switch's founders pursue. Switch will only succeed if this entity is interested in the energy trading platform its founders elect to implement.
6	This entity falls under the governance of the department of public enterprises, and controls a portion of the South African distribution network. It thus possesses considerable structural power. Additionally, this entity possesses some necessity power as it is fundamentally linked to Switch's future development. Finally, this entity possesses considerable status power in its expansive network of connections.	This entity is considered one of the most powerful entities in the South African energy sector. Switch will, for example, not be able to operate in a certain portion of the country's distribution network unless this entity grants it permission to do so (no matter the trading model its founders select).

TABLE 12.1 (continued): *An analysis of the power possessed by the entities deemed affected during Step 3 of the methodology.*

label	Power description	Power Relevance
7	This entity forms part of South Africa's provincial governance structure, and is, furthermore, the official proprietor of a large portion of the country's electricity distribution network. As such, it possesses significant structural power. This entity also possesses considerable necessity power as Switch's future development depends upon its willingness to grant Switch access to the distribution network it controls. Finally, this entity possesses considerable status power in the network of connections it has access to.	This entity is one of the primary gate keepers of Switch's future development. Unless it is interested in the operational value Switch has to offer, no progress will be made toward realising Switch's objectives (no matter the trading model its founders select).
8	This entity possesses status power as a result of the domain-specific knowledge it holds in respect of the South African energy sector, and the network of connections it has access to.	This entity could be approached as an expert in the energy sector in an attempt to verify the validity of Switch's business plan. Additionally, the network of connections it has access to could be utilised as link to potential investors.
9	This entity possesses a degree of incapacitation power. If it is not satisfied with the general changes incurred upon the South African energy sector as a result of the dynamic fee trading strategy Switch incorporates, it may attempt to sabotage the pursuit of Switch's objectives.	Depending upon how the electricity market responds, this entity may sustain losses as a result of Switch's approach to energy trading. If this occurs, it may attempt to inhibit Switch's development. Although this seems unlikely, the possibility should be noted.
10	This entity possesses considerable necessity power as Switch's energy trading software solution cannot be implemented without its participation.	This entity is another gatekeeper to Switch's development. Despite its lack of structural power, the necessity power it possesses provides it with the influence to either seriously deter or facilitate Switch's development.
11	Status power in the network of connections it has access to.	The relevance of this entity's power is unclear. Its network of connections within the energy sector may lead to valuable insights and/or partnerships.
12	This entity possesses considerable necessity power as it controls a third kind of distribution network in which Switch may attempt to establish an energy trading platform.	While the power possessed by this entity is limited in its capacity to inhibit Switch's development, it holds significant potential in facilitating the realisation of the goals Switch's founders pursue.
13	Status power in its operational infrastructure, clientele, brand reputation, and domain-specific knowledge.	This entity may be approached as a source of insight into the difficulties of operating as an energy trading software provider.

TABLE 12.1 (continued): *An analysis of the power possessed by the entities deemed affected during Step 3 of the methodology.*

label	Power description	Power Relevance
14	Status power in the form of capital.	The power possessed by this entity may grant Switch's founders access to capital without having to subdivide the company's equity shares. This will, however, result in the responsibility of having to pay back a loan over a considerable period of time.
15	Status power in the form of capital.	Potential source of capital. Procuring capital form this entity will result in the loss of equity shares.
16	Status power in the form of capital.	Potential source of capital. An advantage of this form of capital funding is that Switch's founders may maintain complete ownership of the company. If done right, furthermore, it has the potential of bringing in a large amount of money.
17	Status power in the form of business support and access to a network of potential investors.	Switch is in the early stages of its development and could benefit from the business development services that this entity has to offer. This will go hand-in-hand with granting it a portion of ownership in Switch. This may, however, be beneficial — Switch will benefit in terms of its brand's reputation, and the network of connections it has access to.
18	Status power in the form of its expertise in the development of blockchain-based software, and its ability to provide business support.	Being a blockchain-specific business incubator, this entity is certainly relevant to Switch's objectives. It would be able to support Switch in the process of developing software, and identifying investment opportunities. It will expect equity shares in return.
19	Status power in the form of business support, the extensive network of investors it has access to and the reputation of its brand.	Switch could benefit from the business development services that this entity has to offer. This will, however, involve granting it equity shares.
20	Status power as a result of its expertise in the energy sector, the network of connections it has access to, the reputation of its brand, its network of clientele, and the capital it has access to.	The power possessed by this entity includes material resources in the form of capital. This entity will expect equity shares in return for capital, but it should be kept in mind that partnering with the entity in question may result in additional benefits such as access to clientele, general business support, and the opportunity to build upon an established operational platform.

TABLE 12.1 (continued): *An analysis of the power possessed by the entities deemed affected during Step 3 of the methodology.*

label	Power description	Power Relevance
21	Status power in the network of connections it has access to, its brand's reputation, and the amount of capital it holds.	This entity could be approached in the capacity of potential investor and/or link to a potential investor. Being invested in by this entity could have considerable brand reputation benefits.
22	Status power in its network of connections and the reputation of its brand.	The network of connections available to this entity could be utilised to identify potential investors.
23	Status power in the network of connections it has access to, its domain-specific knowledge and the reputation of its brand.	This entity could be approached as an expert in the energy sector in order to validate Switch's business plan, or as a link to potential investors.
24	As the official custodian of South Africa's renewable energy sector, this entity possesses considerable structural power. It also possesses status power in the information it holds concerning future developments in the sector in question, and the network of connections it has access to.	Similar to the Council for Scientific and Industrial Research, this entity may be approached as expert in the energy sector in order to determine the validity of Switch's business plan. It may also be able to connect Switch's founders to potential investors.
25	This entity possesses considerable status power as a result of its expertise in, and domain-specific knowledge about, the energy sector, its ability to provide business development support, and the network of connections it has access to.	This entity could be approached as an expert in the energy sector in order to verify the validity of Switch's business plan. It could also be a link to potential investors.
26	Status power in its network of connections (*e.g.* it is endorsed by African Heads of State and Government on Climate Change [23]), and its ability to provide general business support.	This entity may be approached as link to potential investors and/or a source of guidance in respect of Switch's business plan.
27	This entity possesses considerable structural power as its director is second only to the President on matters pertaining to the energy sector [128]. It also possesses considerable status power in the form of capital, the network of connections it has access to, and its ability to sway public opinion.	This entity could be approached in an attempt to draw upon the structural and status power it possesses. It also has the capacity to act as investor. It is worth noting that this entity seems rather difficult to access.
28	This entity possesses considerable status power as a result of its expertise in the energy sector, its network of connections, the reputation of its brand, and the capital it has access to.	This entity may be approached in the capacity of potential investor and/or source of guidance in respect of Switch's business plan. Being invested in by this entity will have considerable brand reputation benefits.

TABLE 12.1 (continued): *An analysis of the power possessed by the entities deemed affected during Step 3 of the methodology.*

label	Power description	Power Relevance
29	This entity possesses considerable status power as a result of the network of connections it has access to, its ability to provide business support, the reputation of its brand, and the capital it holds.	This entity may be approached as potential investor. It provides capital in one of three ways: Grants, loans and venture capital.
30	As director of the Development Bank of Southern Africa, this entity possesses considerable structural power. It also possesses a degree of status power in the network of connections it has access to.	This entity could be approached in an attempt to utilise its influence in order to procure capital funding from the Development Bank of Southern Africa. It is worth noting that the entity in question is a personal acquaintance of one of Switch's founders.

simply as a means to that end, while the Council for Scientific and Industrial Research, PowerX and Energy Exchange, were each selected as stakeholders because of the expert knowledge they possess in respect of matters pertaining to the South African energy sector.

In conclusion to this section, it is worth noting that by meeting with each of the aforementioned entities, Switch's founders were able to identify a viable course of action in respect of Switch's future development. That is, they elected to implement a wheeling-based trading strategy, to do so in the context of municipally-owned distribution networks and privately-owned grid-tied micro-grids, with the assistance of the South African Renewable Energy Business Incubator (in return for an equity share), in partnership with Eldo, and funded by the Development Bank of Southern Africa.

12.3 Applying the decision-making checklist

In §9.3, the decision-making checklist was described as a method designed to facilitate the mitigation of the irrational influence in the mechanisms of human decision making. In order to evaluate the validity of this claim, the method in question was applied to the context of interest in this chapter. More specifically, it was deemed strategically appropriate to subject the decisions made by Switch's founders in respect of its future development to the version of the decision-making checklist presented in §11.2. Notably, in the spirit of this chapter, the subsequent discussion of this evaluation is biased in respect of insight gained.

One of the first insights uncovered during the application of the decision-making checklist concerns the augmentation of Switch's value-offering to grid-tied micro-grid owners. That is, it was determined that Switch could rather easily augment its value-offering by further developing its software application to include functionalities such as water-billing and/or rent-billing, thereby aiding the aforementioned type of entity. In addition to this realisation, the application of the method dedicated to affective failures, resulted in Switch's founders being confronted with the importance of remaining sensitive to how the South African energy market develops in the near future as it was possible that third-party energy trading could become more prevalent. When asked whether this possibility was acceptable, the founders concluded that they had no alterna-

tive but to pursue the selected course of action as third-party energy trading was not yet feasible and that time was of the essence.

The application of the fourth method in the decision-making checklist resulted in two rather interesting discoveries. In respect of the first, recall that one of the core benefits claimed to characterise Switch's value-offering, is that it furthers the vision of a more sustainable South Africa. Although this may be true in general, a technicality was uncovered that may do this claim some harm in the eyes of potential clients situated in municipally-owned distribution networks. In particular, it was uncovered that because of the way in which the South African distribution network functions, *clean* and *dirty* energy cannot be distributed separately (there is only one distribution network). This means that although Switch may be able to facilitate peer-to-peer energy trading from a remunerations point-of-view (*i.e.* a specific producer will be paid by a specific consumer), there is no way of ensuring that the actual energy used by a particular customer is the (clean) energy produced by a particular producer. Ultimately, this was interpreted as another factor suggesting that grid-tied micro-grids was a context better suited to the pursuit of Switch's development[10].

The second insight gained during the application of the aforementioned method concerns the founders' decision to partner with Eldo. The method in question entails re-evaluating any information that contradicts the appropriateness of a selected decision option, and it was during the execution of this task that one of Switch's founders revealed that despite the agreement to partner, Eldo was in the process of developing a product "similar" to its own together with a competitor (*i.e.* Powerx). To the surprise of the author, the founders were initially rather unconcerned by this reality. Based upon further investigation, however, the founders came to agree that the contract between themselves and Eldo had to include precautionary clauses so that they could not be hoodwinked.

The final method in the decision making checklist, dedicated to counteracting the tendency whereby information believed to be contrary to the dominant opinion in a group, revealed two last points of concern. The first was that Switch's founders had yet to confirm the compatibility of their software with the particular kind of smart-meters utilised in municipally-owned distribution networks and/or privately-owned grid-tied micro-grids. The second was that the legalities surrounding the conversion of *power purchasing agreements* into blockchain-based *smart contracts* had yet to be investigated.

To conclude this section, the reader is referred to Appendix B, containing a copy of a review provided by one of Switch's founders concerning the utility of the methods applied to the context of this chapter.

12.4 Chapter summary

This chapter was devoted to an account of the practical application of the meta-methodology proposed in this dissertation in a final attempt to evaluate its utility in facilitating the solution of real-world problems. The chapter opened in §12.1 with a brief introduction to the problem considered, and this was followed in §12.2 and §12.3 with a discussion of actual application of the meta-methodology proposed.

Notably, a quality that was particularly salient during its application, was the importance of a good fit between practitioner and methodology. That is, as a by-product of a dedication to

[10] In the context of a micro grid, prosumers, consumers and the energy they trade remains isolated from the general distribution network.

general applicability, it became evident that there were a few technicalities involved in addressing the problem considered which had not explicitly been accommodated in the meta-methodology (*e.g.* considering the disposition of affected entities in respect of Switch's development). During a meeting with the expert analyst introduced in §2.3, it was determined that despite placing a greater measure of responsibility upon future practitioners, this quality was acceptable as the scope of applicability associated with the meta-methodology was one of its core strengths, one that would be lost if any alteration was made thereto in order to tailor it to a specific kind of problem [463]. As such, the version of the meta-methodology applied is, in fact, the final one proposed.

CHAPTER 13

The mechanisms of interaction

Contents

This chapter is dedicated to an inquiry into the dynamics that govern the mechanisms of human interaction in the context of an endeavour by the mathematical modelling thereof. The inquiry commences with the introduction of a governing assumption in §13.1, followed by an abstract mathematical description of the relationship between the mechanisms of gain and power in §13.2 and §13.3. The effect of infeasibility upon this relationship is then briefly discussed in §13.4, after which an analysis of the proposed model is conducted in §13.5. Next, an extension to the model is presented in §13.6 and §13.7, and this is followed in §13.8 by the generalisation of a result related to the manner in which entities compete in an endeavour, depending on the properties that characterise it. The chapter is finally brought to a close with a short summary in §13.9.

13.1 Rationality as a governing assumption

It is likely impossible to fully capture the extensive range of dynamics relevant to the meta-methodological process introduced in the previous chapter. That does not, however, liberate one from the responsibility of attempting such an exposition. Accordingly, instead of shying away from the challenge, the scope of dynamics to be considered in this chapter is limited to a selection of those involved in the competitive interaction of stakeholders in an endeavour, and furthermore, are studied under highly idealised conditions. Before proceeding with such an exposition, it is acknowledged that the model presented in this chapter is not a tool designed to facilitate real-world optimisation; instead its purpose is twofold: To present the reader with a

manner in which to conceptualise the context of an endeavour, and to venture to uncover certain dynamics related thereto which may otherwise be hidden from a purely contemplative analysis[1].

It should be evident from the literature review of Chapters 3 and 4 that the dynamics which dominate stakeholder interaction relate to either the *cognitive* or *social* mechanisms of the *self*. In order to discuss these mechanisms, and subsequently model them, it is necessary to employ the following governing assumption:

A1. *The nature of decision making.* Whenever an entity is presented with several decision options, its intention will be to employ its cognitive capacity in a manner that an objective party would consider rational — that instances of poor decision making are examples of occasions on which entities were unable to employ the mechanisms of human decision making appropriately, and that such cases do not represent the norm, but the exception.

This assumption is the first of several that escort this work from what constitutes reality, to what one may consider highly idealised conditions. Nonetheless, to illustrate that the above assumption is not entirely unrealistic, the cognitive and social mechanisms related thereto are discussed in slightly more detail. Or rather, the most common instances in which Assumption A1 does not hold are discussed to illustrate that such cases are not entirely ubiquitous. Fortunately, these cases have, in fact, already been described. Recall the notion of a cognitive error defined in §9.3 to represent the case in which the cognitive mechanisms of the mind fail to respond appropriately to a presented decision option. And similarly, recall the notion of a social error defined in §9.3 to represent the case in which certain qualities in an entity's environment cause a misalignment between the outcomes it would pursue if it were not influenced by its environment, and the temporary context-imposed outcomes it pursues. It follows that in order to defend Assumption A1 it is necessary to illustrate that the cases of the cognitive and the social error are, in fact, the exception rather than the rule. In terms of the cognitive error this should, however, be rather obvious as the entirety of Chapter 3 was dedicated to an explication of just that fact. Consequently, despite the reality of cognitive errors, the governing assumption can be considered to remain intact in that respect. The case of the social error is somewhat more difficult to defend.

As described in §4.2.2, the (occasionally) variable nature of the self can be traced back to the notion of a working self-concept. More specifically, a contextually activated working self-concept can influence an entity's inner-working so that the inner variability of its social state manifests itself in the form of irregular behaviour [105]. In such a case the foundation from which decisions are being made is temporarily misaligned with the objectives the entity would have pursued if it were not influenced by the social dynamics of the environment. Toward understanding the true nature of the variability that exists in the real-time nature of the self, the importance of consistency in behaviour, as noted in §4.1, deserves repeating. Importantly, the variability in the social state of the self does not operate in an unrestrained manner, but co-exists with other conflicting influences, one of which is the inherent value the human psyche places upon consistency in behaviour. Moreover, it should be noted that an individual's self-concept is not contained within only its social component. As defined in §4.2, the universal self consists of three components, namely the *individual self*, the *relational self*, and the *collective self*. Significantly, the changing nature of the working self-concept is contained within the mechanisms of the relational and collective selves. In other words, the individual self represents the part of the universal self that values consistency in behaviour and stands in conflict with the changing nature of the collective and relational selves. The question of interest is thus which of these selves can be described as dominant? Although there is not complete agreement in the academic literature

[1] Such findings should be interpreted with respect to the idealised conditions under which they are uncovered.

in this respect, the majority of researchers hold that the hierarchy among these elements is as follows: First the individual self, then the relational self, and finally the collective self [408]. In other words, consistency in behaviour dominates variability and consequently, Assumption A1 is appropriate.

In summary, although the discourse in this section is in no way a denial of the fact that human perception is vulnerable to irrational influences — the entirety of this dissertation, and more specifically, the two cases discussed above, acknowledges this potential for irrational decision making. What is, however, assumed is threefold: That entities are never intentionally irrational, that instances of irrational decision making are the exception rather than the rule, and similarly, that instances of rational, but poor, decision making are also the exception rather than the rule. Consequently, the intent of Assumption A1 is not to contradict what has been said before — it is to argue that individuals will attempt to make decisions in such a way that harmony exists between their understanding of a problematic situation, their perception of the consequences of a specific decision option, the outcomes they are pursuing and the actual decision that is made, and furthermore, that they will attempt to do this in a rational fashion.

13.2 The tracery of rationality

When faced with several decision options, it is considered rational to note one's pursued outcomes, relate those outcomes to one's ability in realising them and proceed to select a decision option that maximises one's expected gain within the overlap of gain and ability [134]. This description of rational choice illustrates the central role played by the ability to realise pursued outcomes and the expected gain of those outcomes in the selection of a rational decision option. In order to elicit dynamics related to the interaction between several competing entities, a good place to start is thus to describe the two elements that tend to dominate each entity's individual (assumed rational) decision making process (*i.e.* an entity's realisation ability and the expected gain associated with the outcomes pursued). More specifically, the question of interest is whether one can mathematically describe the relationship between an entity's level of realisation ability and the measure of gain it can expect with respect to outcomes pursued and, furthermore, whether such a description may be leveraged to uncover intricacies related to the competitive interaction of individuals that may be hidden from a purely contemplative analysis.

What makes such an exposition difficult is the fact that the nature of the elements constituting realisation ability and expected gain typically depend on the system in which they are described. Consequently, the notions by which these elements are represented have to be tailored to the specific kind of system in which they are employed. It is thus necessary to define two tailor-made elements by which realisation ability and expected gain can be represented specific to the context of an endeavour. Before embarking on such an exposition, however, there are several assumptions, elements, and properties pertaining to the working of an endeavour that require clarification. First and foremost, the following assumption should be noted with respect to the model proposed in this chapter:

A2. *The theoretical nature of what follows.* The current investigation into the dynamics of realisation ability and expected gain is conducted in a theoretical capacity and, as such, the elements relevant to it are assumed fully known, and thus, accurately describable.

Secondly, it is necessary to define the notion of an *entity* formally. An entity represents a single or collective group of individuals working toward the realisation of a collective set of outcomes within the context of an endeavour. In other words, the pursuit of a single outcome not shared

distinguishes one entity from another. Another notion in need of definition is that of a *pursued outcome*. In the context of an endeavour, a pursued outcome simply represents a certain objective i pursued by a certain entity x. For every entity x it is therefore possible to describe a set of objectives $\mathcal{I}_x$ that represents the outcomes pursued by that entity. It follows, furthermore, that one can define a complete set of objectives $\mathcal{I} = \cup_x \mathcal{I}_x$ representing the collection of generally pursued outcomes within an endeavour. Importantly, among any two objectives[2] $i, j \in \mathcal{I}$, there may exist a degree of correlation — a notion not be confused with the practical reality of entities working together in the spirit of collaboration. The following assumption is made with respect to the modelling of entities:

A3. *The competitive entity.* The interaction among the entities of an endeavour is assumed to be fundamentally competitive — entities are considered to value only the attainment of their own objectives.

Consequently, the notion of objective correlation cannot be likened to a group of investors aggregating their capital for mutual benefit[3]. It simply represents the unintended overlap of the practical outcomes represented by two or more objectives. With these dynamics in mind, the elements representative of realisation ability and expected gain may now be introduced.

Toward describing the mechanisms of expected gain, recall the notion of effect, defined in §9.2.1 to represent the degree to which an entity is influenced by the outcome of an endeavour. There are several points of similarity between this notion, and what expected gain has been designed to represent. More specifically, the measure of gain of an entity x expected in the pursuit of an objective $i \in \mathcal{I}_x$, denoted by G_x^i, is representative of any form of benefit anticipated in the pursuit of that objective in the context of the same functional network employed to define the notion of effect in §9.2.1. The gain expected in the pursuit of an objective i is thus representative of the positive component of effect pertaining to the pursuit of i.

In like manner, the description of realisation ability also employs an element previously defined. The notion of *power*, defined in §9.2.1 to represent an entity's capacity to influence the decisions made by another entity, and the ability to realise outcomes pursued, represent the same dynamic. As such, the elements that constitute the notion of power represent the mechanisms by which outcomes pursued are realised within the context of an endeavour. In other words, the ability of entity x to realise pursued outcomes can be described in terms of power possessed (*i.e.* assuming that the measure of an entity's power can, in fact, be determined according to some scale). Furthermore, because the measure of entities' power may differ for different objectives, it is beneficial to describe its measure with respect to a specific objective $i \in \mathcal{I}$. Accordingly, the realisation ability of an entity x in the pursuit of an objective i is described in terms of its measure of power possessed in the pursuit of objective i, denoted by $\tilde{P}_x^i$. It follows that the notion of this power links back the dimension defined to constitute the notion of power in §9.2.1.

As one cannot compare apples with pears, the issue of unit conversion must be addressed before relating the mechanisms of power to that of expected gain. To that end, $\tilde{P}_x^i$ is normalised in order to create a unitless representation of its measure. Consequently,

$$P_x^i = \frac{\tilde{P}_x^i}{\max\limits_{i,x} \tilde{P}_x^i},$$

[2] The label assigned here to a specific objective refers to its membership of the set of generally pursued objectives $\mathcal{I}$, and not the entity-specific set of pursued objectives $\mathcal{I}_x$.

[3] Such a group of investors would be considered a single entity.

where P_x^i represents the unitless measure of power contributed toward the pursuit of an objective $i \in \mathcal{I}_x$ by an entity x in the unit interval $[0, 1]$. An important point should be noted here. Although it is possible to describe the relationship between expected gain and possessed power in general in terms of P_x^i, it is deemed unwise to do so at this point in the discussion as such an analysis would be exceedingly difficult to carry out. A more appropriate strategy is first to introduce perhaps the simplest mathematical form describing the relationship between possessed power and expected gain, and then proceed to increase the complexity thereof incrementally. To that end, a more general element — the total measure of *effective power* associated with the pursuit of a particular objective $i \in \mathcal{I}$ by the entities partaking in the endeavour is denoted by P^i. In other words, while P_x^i quantifies the power contributed by a single entity x in the pursuit of a certain objective $i \in \mathcal{I}_x$, P^i quantifies the total power effective in the pursuit of i (*i.e.* not specific to a single entity). As will be illustrated subsequently, P^i depends on several unitless parameters, and it too is unitless.

The notion of power has been described relative to a specific objective, and so it is necessary to apply a similar modification to the notion of gain expected. Fortunately, however, because of the nature of the elements that constitute it, an endeavour possesses the rather interesting property that the measure of gain expected by an entity x in the pursuit of an objective $i \in \mathcal{I}_x$ is equal to the combined measure of gain G^i expected in the general pursuit of that objective by all entities x associated with the endeavour for which $i \in \mathcal{I}_x$. As such, although counter-intuitive, the variables G_x^i and G^i are essentially equal for all x. To elucidate this claim, objectives by and large belong to only a single entity because of the competitive nature of an endeavour. The only case in which entities are able to share in the pursuit of an objective is if they exhibit the property that the gain attained in the pursuit of that objective can be shared without being divided. Noting that $G_x^i = 0$ for all objectives $i \notin \mathcal{I}_x$, it follows that $G^i = G_x^i$ for any x because $\sum_x G_x^i = G_x^i$ for any objective $i \in \mathcal{I}$, or because of the property above.

In order to describe the relationship between expected gain and power, the variables G^i and P^i are employed. Note the effect of substituting the notion of realisation ability for that of power, and the notion of an outcome for that of an objective in the description of rational choice with which this section opened:

> When faced with several decision options, it is considered rational to note one's pursued objective(s), relate those objectives to one's measure of power possessed (with respect to pursuing the objectives in question) and proceed to select a decision option that maximises one's expected gain within the overlap of gain and power.

It should be apparent from this description of rational choice that the measure of gain expected in an entity's pursuit of an objective depends upon the general measure of effective power associated with the pursuit of that objective. Because this property holds for each objective $i \in \mathcal{I}$, the gain expected in pursuit of any objective i is proportional to the measure of effective power generally associated with its pursuit. In other words, noting that G^i and P^i represent these respective measures, it follows that

$$G^i \propto P^i, \qquad i \in \mathcal{I}. \tag{13.1}$$

The expected gain G^i is thus an increasing function of P^i, such that the pursuit of objectives within an endeavour can be likened to a balancing act between power possessed and gain expected — a fact that is certainly insightful, but to truly draw benefit from expression (13.1), it is necessary to develop it further. Before doing so, however, another point should be made in respect of the kind of system for which the model proposed in this chapter is relevant. It should

be apparent that the dynamics related to an endeavour that enters, and remains in, a state of chaos will not follow the expression in (13.1). As such, the arguments that follow are relevant only to endeavours that conclude in a state of order[4].

13.3 The mechanisms of power

The pursuit of an objective was described in the previous section as a balancing act between power possessed and gain expected. As power constitutes the independent component of these elements, it is appropriate to proceed with its explication. The measure of effective power P^i associated with the pursuit of an objective $i \in \mathcal{I}$ contains two dimensions: The total measure of power working toward its realisation, denoted by T^i, and the total measure of power working against its realisation, denoted by C^i. In order to define these elements more formally, several assumptions related to the pursuit of objectives in the context of an endeavour should be noted. The first of these assumptions addresses the nature of the variable P_x^i, defined in the previous section to represent the measure of power possessed by an entity x with respect to the pursuit of objective i. The following assumption is made in order to model the competitive pursuit of objectives within an endeavour.

A4. *The nature of power possessed.* It is assumed that the measure of power P_x^i of an entity x with respect to the pursuit of an objective i contains only the portion of its power that works toward maximising the attainment of i.

As such, the only way in which an entity can reduce the attainment of an objective, is if it pursues the realisation of another objective negatively correlated to it. The degree of correlation between two objectives $i, j \in \mathcal{I}$ is denoted by the parameter $\delta^{ij} \in [-1, 1]$, where a positive value of δ^{ij} represents positive correlation between the objectives, and a negative value of δ^{ij} represents negative correlation between objectives. Now consider an entity x working toward the realisation of the objectives in $\mathcal{I}_x$. It should be apparent that if provided with such information, an entity would concern itself only with the objectives in the set $\mathcal{I} \setminus \mathcal{I}_x$ that are actually capable of influencing the pursuit of the objectives in $\mathcal{I}_x$ (*i.e.* objectives in $\mathcal{I} \setminus \mathcal{I}_x$ that share a degree of correlation with the objectives in $\mathcal{I}_x$). Furthermore, based upon Assumption A1, it should be apparent that an entity x would typically not participate in the pursuit of any objective $j \in \mathcal{I} \setminus \mathcal{I}_x$ for which $\delta^{ij} \leq 0$, for any $i \in \mathcal{I}_x$, but that it may, if possible, want to participate in the pursuit of an objective $j \in \mathcal{I} \setminus \mathcal{I}_x$ for which $\delta^{ij} > 0$, for any $i \in \mathcal{I}_x$, if doing so furthers the attainment of $i \in \mathcal{I}_x$. Upon defining the set $\mathcal{I}_x^+$ to represent the union of the set $\mathcal{I}_x$ and the set of objectives $j \in \mathcal{I} \setminus \mathcal{I}_x$ for which $\delta^{ij} > 0$, for some $i \in \mathcal{I}_x$, the following assumption is thus made:

A5. *The case of bounded interest.* It is assumed that an entity x will participate only in the realisation of the select group of objectives in $\mathcal{I}_x^+$. In other words, the measure of power P_x^i possessed by an entity x with respect to the pursuit of an objective i, will only influence the attainment of $i \in \mathcal{I}$ directly (*i.e.* not through the medium of objective correlation) if it also belongs to the set $\mathcal{I}_x^+$.

It follows from Assumptions A4 and A5 that the magnitude of T^i depends on the measure of power possessed by entities working toward the realisation of objective i, or some objective $j \in \mathcal{I}$

[4]This assumption does not preclude the potential for chaos to emerge at some point during an endeavour. Irrespective of what takes place during an endeavour, the discussion presented in this chapter is relevant only to endeavours that conclude in a state of order.

positively correlated with i (*i.e.* if the attainment of objective j increases, that of objective i will increase accordingly). And similarly, the magnitude of C^i depends on the measure of power possessed by entities pursuing the realisation of some objective $j \in \mathcal{I}$ which is negatively correlated with i (*i.e.* if the attainment of objective j increases, that of objective i will decrease accordingly). Before attempting to describe these elements mathematically, it is necessary to clarify a technicality about the manner in which the measure of an entity's power is affected by its employment.

Power was defined in §9.2.1 to exist along four dimensions — form, source, efficacy and potential. Recall that the dimension labelled source, in fact, contains several formal and informal power sources. In a practical context, it should be apparent that the employment of such a power source may result in the reduction of the level of power associated with it (*i.e.* its potential). More specifically, if an entity x employs its power P_x^i towards the realisation of an objective $i \in \mathcal{I}_x^+$, the power P_x^j possessed by that entity, relevant to another objective $j \in \mathcal{I}_x^+$, may be affected. Despite the reality of this dynamic, the following assumption is made in order to model the competitive interaction among the entities of an endeavour:

A6. *Independence of power-plays.* The employment of the power P_x^i of an entity x with respect to the pursuit of an objective $i \in \mathcal{I}_x^+$ does not alter that entity's power P_x^j with respect to the pursuit of another objective $j \in \mathcal{I}_x^+$, for all $i, j \in \mathcal{I}_x^+$.

The above assumption may seem to be a serious oversimplification of the real world. It should be noted, however, that three of the five sources of power discussed in §9.2.1 do, in fact, possess the unique quality that their potential does not diminish with use. More specifically, the only elements vulnerable to variation as a result of employment are contained in the power sources of status and incapacitation (and only when they are employed in a coercive fashion). Consequently, although a simplification, Assumption A6 is more descriptive of the modelled system than not. With this assumption in mind, two additional variables can now be defined formally. The power T^i working *toward* the realisation of an objective i can be represented by

$$T^i = \sum_{x=1}^{m} \sum_{j \in \mathcal{I}_x^+} \delta^{ij} P_x^j, \qquad i \in \mathcal{I}. \tag{13.2}$$

Similarly, the measure of power C^i working *against* the realisation of an objective i can be represented by

$$C^i = \left| \sum_{x=1}^{m} \sum_{j \in \mathcal{I} \setminus \mathcal{I}_x^+} \delta^{ij} P_x^j \right|, \qquad i \in \mathcal{I}. \tag{13.3}$$

In (13.2) and (13.3), the reason for summing over only $j \in \mathcal{I}_x^+$ and $j \in \mathcal{I} \setminus \mathcal{I}_x^+$, respectively, lies in Assumption A5. The nature of the elements introduced above is more significant than it may at first appear. To elicit this significance, it is beneficial to investigate the intricacies of objective correlation in more detail by means of a hypothetical example.

Example 13.1. *Consider the following illustrative example of the nature of objective correlation. Figure 13.1 depicts the correlation of a base-objective A together with three objectives B, C and D, pursued by four entities a, b, c and d, respectively, according to the following conventions: First, objectives are modelled as unit-vectors in order to facilitate calculation of the correlation of one objective with another. Secondly, the degree of correlation δ^{ij} between two objectives i and j is represented by the counter-clockwise angle $\theta^{Aj} \in [0, \pi]$ between them, where $\theta^{Aj} = \cos^{-1} \delta^{Aj}$ for $j \in \{B, C, D\}$ (i.e. if $\theta^{Aj} < \frac{\pi}{2}$, then objectives A and j are positively correlated, while if $\theta^{Aj} > \frac{\pi}{2}$, the objectives A and j are negatively correlated).*

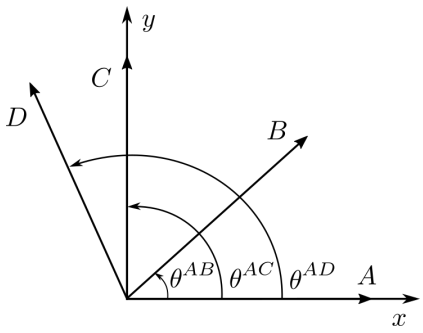


FIGURE 13.1: *The alignment of objectives with the base-objective, labelled A.*

According to the conventions introduced above, it follows that the degree of correlation between objective A and any other objective $j \in \{B, C, D\}$ is the vector component of A in the direction of j, and vice-versa. The degree of correlation between the objectives A and B in Figure 13.1, can thus be expressed as $\cos\theta^{AB}$, while $\sin\theta^{AB}$ represents the magnitude of the component of objective B independent of A, as depicted in Figure 13.2.

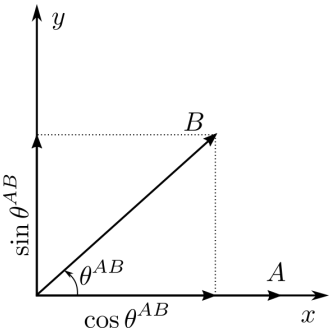


FIGURE 13.2: *The vector components of objective B in the direction of objective A and perpendicular to it, where $\cos\theta^{AB}$ represents the degree of correlation between objectives A and B.*

By determining the respective vector components of objectives B, C, D in the direction of objective A, the interaction between objectives A, B, C and D can be captured with respect to only the components of objectives B, C and D that are actually relevant to the pursuit of objective A, as depicted in Figure 13.3.

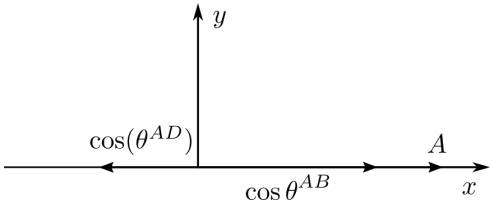


FIGURE 13.3: *The vector components of objectives B, C and D that are relevant to the pursuit of objective A.*

The practice of neglecting the components of objectives that are independent of one another, is useful when modelling the interaction between the entities of an endeavour. To illustrate this,

note that the components of objectives A, B, C and D in Figure 13.3 are, in fact, fundamental to the measure of T^A *and* C^A*. More specifically,*

$$T^A = \cos\theta^{AA} P_a^A + \cos\theta^{AB} P_a^B + \cos\theta^{AB} P_b^B + \cos\theta^{AA} P_b^A,$$

where P_a^A *represents the power possessed by entity a in the pursuit of objective A,* $\cos\theta^{AA} = 1$ *represents the degree of correlation between objectives A and A,* P_a^B *represents the power possessed by entity a in the pursuit of objective B,* $\cos\theta^{AB}$ *represents the degree of correlation between objectives A and B,* P_b^B *represents the power possessed by entity b in the pursuit of objective B, and* P_b^A *represents the power possessed by entity b in the pursuit of objective A. Similarly,*

$$C^A = \cos\theta^{AD} P_d^D,$$

where P_d^D *represents the power possessed by entity d in the pursuit of objective D, and* $\cos\theta^{AD}$ *represents the degree of correlation between objectives A and D. Note that the reason for not considering*[5] *the power of entities a and b with respect to objective D, resides in Assumption A5. Like the elements that constitute them, the scalars* T^A *and* C^A *are unitless and, furthermore, the influence of all objectives relevant to the pursuit of A are contained in these quantities.* □

The fact that T^i and C^i contain the influence of all objectives $j \in \mathcal{I}$ relevant to the pursuit of objective $i \in \mathcal{I}_x$ make it possible to describe the effective measure of power P^i as a function of T^i and C^i. In order to derive the form of the function $P^i(T^i, C^i)$ it is necessary to elicit the practical dynamics one would expect from the interaction between P^i, T^i and C^i. As described in §9.2.1, stakeholders relevant to an endeavour are selected either because of their capacity to affect the outcome of that endeavour, or because of the degree to which they are affected thereby. In other words, if an entity is selected as a stakeholder, the concerns of that entity are deemed both significant and relevant. As such, the following assumption is made regarding the system being modelled:

B1. *The resilience of power.* The measure of power working toward the realisation of an objective can never be made entirely redundant by the measure of power working against its realisation. Essentially, the only way an objective can be excluded from consideration is if the effective measure of power P^i associated with its pursuit is zero, and as P^i is determined according to the degree to which the value of T^i compares with that of C^i, the only way that can happen is if the total measure of power T^i working toward its realisation is zero.

As eluded to in Assumption B1, the effective measure of power P^i associated with the pursuit of an objective i depends on the partition of T^i into two components (according to the comparative magnitudes of T^i and C^i): One component representing the portion of T^i that is effective in the pursuit of objective i, and the other representing the portion of T^i that is ineffective in the pursuit of objective i, as a result of C^i. According to the definition of P^i, it follows that P^i is the component of T^i effective in the pursuit of objective i. This can be modelled by the assumption that $P^i \propto T^i$, where the non-constant coefficient of proportionality is itself a proportion describing the manner in which the quantities T^i and C^i compare. Obviously, there are many functional forms that may adequately represent this non-constant coefficient of

[5] Objective C has been neglected from consideration as it plays no part in the pursuit of objective A.

proportionality, while simultaneously satisfying Assumption B1. The simplest of these functional forms is arguably the proportion $T^i/(T^i + C^i)$. Employing this proportion, the exemplar

$$P^i = \left(\frac{T^i}{T^i + C^i} \right) \times T^i \tag{13.4}$$

of the effective power P^i associated with the pursuit of an objective is proposed. The expression in (13.4) captures the dynamics expected from the interaction between P^i, T^i and C^i adequately and, as such, is deemed an appropriate example of a form descriptive of the function $P^i(T^i, C^i)$. Furthermore, for the sake of an argument subsequently made, note the following assumption about the nature of the variables T^i and C^i in (13.4).

B2. *The unconstrained variation of power.* The power T^i working toward the realisation of an objective i can be altered without influencing the power C^i working against its realisation.

In view of the assumption above, the expression in (13.4) can now be employed to describe the relationship between the gain expected in, and the power associated with, the pursuit of an objective $i \in \mathcal{I}$ in a more insightful manner. It follows from (13.1) that the gain expected in the pursuit of an objective i may be expressed as an increasing function of the manner in which T^i and C^i related to that objective compare, as expressed in (13.4), such that G^i is a function of $P^i(T^i, C^i)$.

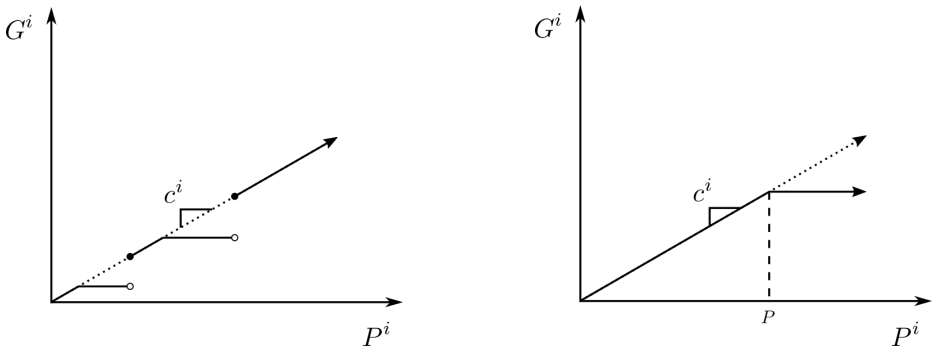
Together, the expressions presented above describe mathematically the manner in which the notion of power may translate to that of gain in the context of an endeavour. The logical next step would be to subject these expressions to an in-depth analysis toward eliciting the dynamics hidden in the mechanisms of their working. Before conducting such an analysis, however, there is a final phenomenon that has to be considered — the *notion of infeasibility*.

13.4 The notion of infeasibility

The notion of infeasibility has been neglected from the preceding theoretical exposition on the relationship between expected gain and possessed power, as it exists only in a practical capacity. But, as this discussion is intended for the practically minded thinker, its influence is briefly addressed in this section. Not surprisingly, the possibility of infeasible pursuits in a practical context has a constraining effect on the relationship between power and expected gain, so that an entity may be unable to attain the full measure of gain merited by the power it possesses.

In essence, the notion of infeasibility causes expected gain to be non-proportional to power for certain measures of power. More specifically, the possibility of infeasibility inhibits the attainment of gain so that the effective measure of power P^i associated with the pursuit of an objective i may increase without an associated increase in expected gain G^i. Uncovering an exhaustive set of such cases is not possible in general. It is sufficient simply to take note of the kind of effect infeasibility may have upon the relationship between power and gain. To that end, consider the two instances of infeasibility presented in Figure 13.4. Note that for the purpose of illustration, G^i is briefly assumed to be *directly* proportional to P^i.

Figures 13.4(a) and 13.4(b) depict two cases in which the assumption $P^i \propto G^i$ does not hold. Figure 13.4(a) represents the case in which no practical solution exists on the line $G^i = c^i P^i$ for certain measures of power, called the *discrete choice problem*. Similarly, Figure 13.4(b) represents the case in which an increase in power is prevented externally from facilitating an increase of expected gain, for some $P^i \in [p, \infty)$, called the *limited resource problem*.



(a) *Discrete choice example* (b) *Limited resource example*

FIGURE 13.4: *Two examples of instances in which infeasibility affects the relationship between power and expected gain.*

13.5 An analysis of the mathematical model

In general, the development of a mathematical model is only as valuable as the hidden complexities of the modelled phenomena it is able to unveil. The model presented in §13.1–§13.3, derived to capture (to an extent) the interaction among the entities of an endeavour, is no different. Accordingly, this section is devoted to an analysis of the model with the purpose of uncovering such hidden intricacies. More specifically, the subsequent analysis is biased toward elements relevant to the balance that exists between the notion of expected gain, and that of power.

Recall that (13.4) represents the total measure of effective power P^i associated with the pursuit of a single objective i. In view of the aforementioned bias, the elements that constitute this notion, and more specifically, the manner in which they affect its value, is of particular interest. Toward uncovering these dynamics, consider the surface plot of the function $P^i(T^i, C^i)$ exemplar in (13.4) for $T^i, C^i \in [0, 1]$ presented in Figure 13.5.

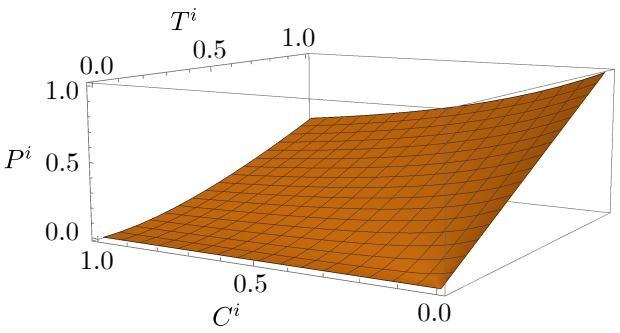


FIGURE 13.5: *A surface plot of the function $P^i(T^i, C^i)$ in (13.4) on the unit square $[0, 1] \times [0, 1]$.*

For the most part, the depicted value of P^i for $T^i, C^i \in [0, 1]$ coincides with what intuition would suggest. The maximum value of P^i, for example, is attained when $C^i = 0$ and $T^i = 1$, and similarly, the minimum value of P^i is attained on the line $T^i = 0$. Significantly, however, Figure 13.5 also uncovers a rather surprising dynamic — it would seem that P^i is more sensitive to changes in T^i than to changes in C^i everywhere on the unit square $[0, 1] \times [0, 1]$. The truth of this conjecture may be verified by comparing the absolute values of the partial derivatives of

the function $P^i(T^i, C^i)$ with respect to T^i and C^i. These partial derivatives of $P^i(T^i, C^i)$ are

$$\frac{\partial P^i}{\partial T^i} = \frac{T^i(2C^i + T^i)}{(C^i + T^i)^2} \tag{13.5}$$

and

$$\frac{\partial P^i}{\partial C^i} = -\frac{(T^i)^2}{(C^i + T^i)^2}. \tag{13.6}$$

Note that

$$\begin{aligned} \frac{\partial P^i}{\partial T^i} &= \frac{T^i(2C^i + T^i)}{(C^i + T^i)^2} \\ &= \frac{(T^i)^2}{(C^i + T^i)^2} + \frac{2T^iC^i}{(C^i + T^i)^2} \\ &= \left|\frac{\partial P^i}{\partial C^i}\right| + \frac{2T^iC^i}{(C^i + T^i)^2}. \end{aligned}$$

In other words, $\partial P^i/\partial T^i \geq |\partial P^i/\partial C^i|$ for all $T^i \in [0, 1]$ and $C^i \in [0, 1]$ — a result that can also be seen in the surface plot of the partial derivatives in (13.5) and (13.6) for $T^i, C^i \in [0, 1]$ in Figure 13.6.

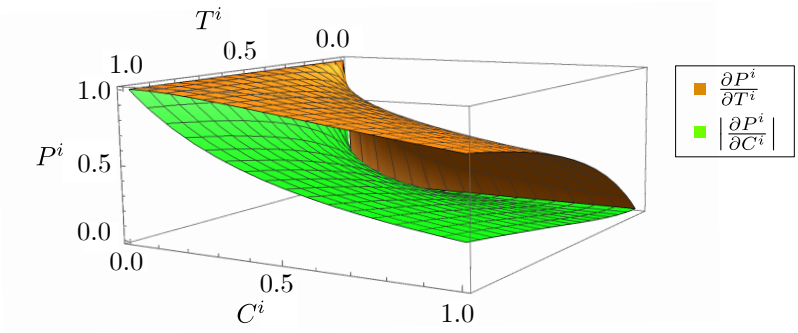


FIGURE 13.6: *A surface plot of the partial derivatives of $P^i(T^i, C^i)$ in (13.5)–(13.6) with respect to T^i and C^i on the domain $[0, 1] \times [0, 1]$.*

The fact that P^i in (13.4) is more sensitive to changes in T^i than to changes in C^i for all $T^i, C^i \geq 0$ is an interesting result. It implies that within any endeavour for which the expression in (13.4) is relevant, an entity would benefit more from increasing the power working toward the realisation of its objectives by a certain value, than it would from attempting to decrease the power working against the realisation of its objectives by the same value. From the expression in (13.2) it is clear that an entity can accomplish such a change in the value of T^i in one of two ways: By increasing the degree of correlation between objective i and any objective $j \in \mathcal{I}$ for which the $\delta^{ij} > 0$, or by increasing the measure of its own power P_x^j with respect to any objective $j \in \mathcal{I}_x^+$. Practically, this could encourage entities to attempt to define their objectives in such a manner that the degree of their correlation is maximised, while simultaneously preserving the practical outcomes pursued.

The potential significance of the above result begs the question of its general applicability — is P^i more sensitive to changes in T^i for *all* endeavours satisfying Assumptions B1–B2, or only for those in which the expression (13.4) is relevant? Before attempting to demonstrate the

generality of this phenomenon, there are several general properties of the function $P^i(T^i, C^i)$ that merit clarification. Note the general[6] boundary conditions

$$P^i(0, C^i) = 0, \qquad C^i \geq 0, \tag{13.7}$$

$$P^i(T^i, 0) = T^i, \qquad T^i \geq 0, \tag{13.8}$$

imposed on the function $P^i(T^i, C^i)$, illustrated graphically in Figure 13.7. These boundary conditions are appropriate as the power P^i associated with the pursuit of an objective i represents the component of T^i that is effective in the pursuit of objective i, which, in turn, depends upon the manner in which the value of T^i compares to that of C^i, as stated in Assumption B1.

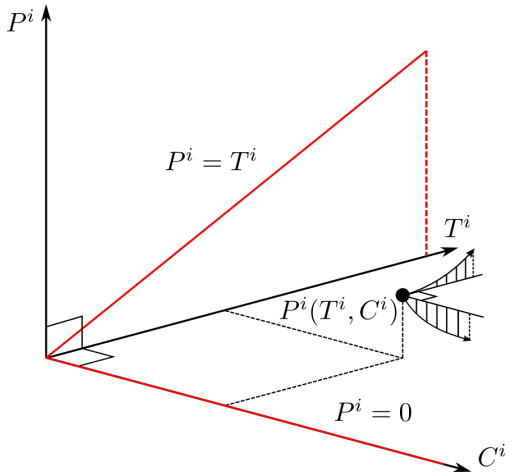


FIGURE 13.7: *Boundary conditions imposed on the function* $P^i(T^i, C^i)$.

With these boundary conditions in mind, note that Figure 13.7 depicts another dynamic in need of clarification — the growth behaviour of the function P^i at a general point (T^i, C^i), more formally illustrated in Figure 13.8. As stated, P^i represents the component of T^i that is effective in the pursuit of objective i. Noting that this effective/ineffective partitioning of T^i depends upon the manner in which the value of T^i compares with that of C^i, it follows that the greater the value of T^i, the greater the effective component of T^i, and thus, the greater the value of P^i. Consequently, as depicted in Figure 13.8(a), the rate of change of P^i as T^i increases, for a constant C^i, is strictly greater than zero. Note, furthermore, that the second-order rate of change of P^i as a function of increasing T^i, for a constant C^i, is strictly non-negative. This is because the second-order rate of change of P^i as a function of increasing T^i, for a constant C^i, essentially represents the extent to which the magnitude of P^i changes as a result of changing T^i, as T^i increases and C^i is held constant. It follows that as T^i increases, for a constant C^i, the effect that C^i has upon the magnitude of P^i will become less and less, but never negligible, and conversely, the effect of T^i upon P^i will continually increase. In other words, as depicted in Figure 13.8(a), the second-order rate of change of P^i as a function of increasing T^i, for a constant C^i, is strictly non-negative. For C^i equal to some constant c, furthermore, the function $P^i(T^i, c)$ is convex and bounded from below by the line $P^i = T^i + q(c)$, also depicted in Figure 13.8(a).

Consider Figure 13.8(b), depicting the growth behaviour of P^i as a function of increasing C^i, for a constant T^i. As can be seen, the rate of change in P^i as a function of increasing C^i, for a constant T^i, is strictly negative. The reason for this is once again related to the effective/ineffective

[6] Although these boundary conditions are certainly satisfied by the specific function P^i in (13.4), it is argued here that the boundary conditions hold in general as a direct consequence of Assumptions B1–B2.

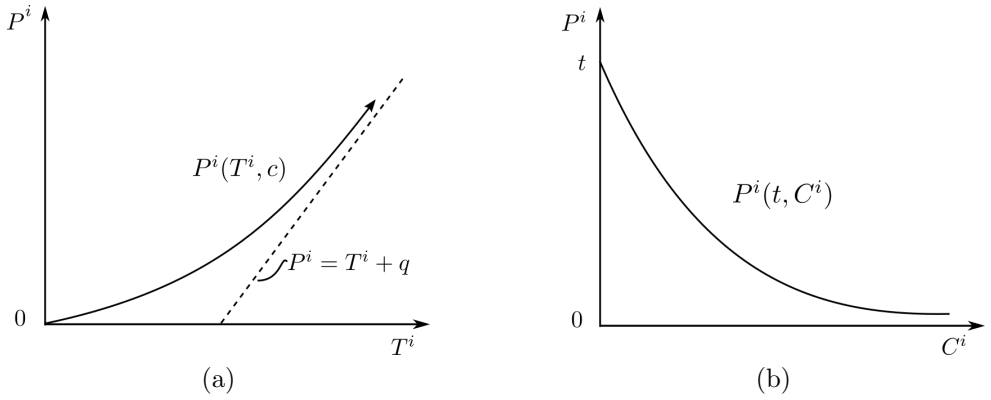


FIGURE 13.8: *(a) The behaviour of $P^i(T^i, C^i)$ as a function of increasing T^i for a fixed value $C^i = c$, and (b) the behaviour of $P^i(T^i, C^i)$ as a function of increasing C^i for a fixed value $T^i = t$.*

partitioning of T^i. It follows that the greater the magnitude of C^i, the smaller the effective portion of T^i, and thus, the smaller the magnitude of P^i. Figure 13.8(b) furthermore illustrates that the second-order rate of change of P^i as a function of increasing C^i, for a constant T^i, is strictly non-negative. This is once again because the second-order rate of change of P^i as a function of increasing C^i, for a constant T^i, represents the extent to which the magnitude of P^i changes as a result of changing C^i, as C^i increases and T^i is held constant. It follows that as C^i increases, for a constant T^i, the effect T^i has upon the magnitude of P^i will become less and less, but never negligible, and conversely, the effect of C^i upon P^i will continually increase. Consequently, for T^i equal to some constant t, the function $P^i(t, C^i)$ is convex for all $C^i \geq 0$. The dynamics discussed in this, and the previous paragraph, relating to the growth behaviour of the function $P^i(T^i, C^i)$, can be described mathematically as

$$\frac{\partial P^i}{\partial T^i} > 0 \quad \text{and} \quad \frac{\partial^2 P^i}{\partial (T^i)^2} \geq 0 \quad \text{for all} \quad (T^i, C^i) \in \mathbb{R}_2^+ \tag{13.9}$$

and

$$\frac{\partial P^i}{\partial C^i} < 0 \quad \text{and} \quad \frac{\partial^2 P^i}{\partial (C^i)^2} \geq 0 \quad \text{for all} \quad (T^i, C^i) \in \mathbb{R}_2^+. \tag{13.10}$$

There is one more technicality to describe before attempting to prove the general applicability of the above result. It is necessary to illustrate the behaviour of the partial derivatives of $P^i(T^i, C^i)$ at $C^i = 0$. It follows from the boundary condition described in (13.8) that an increase in the value of T^i by some constant w, will result in an increase of $P^i(T^i, 0)$ by the same value w. Similarly, a decrease in the value of T^i by some constant w, will result in the decrease of $P^i(T^i, 0)$ by the same value w. Furthermore, based on the Assumption B1, the effect of C^i upon the value of P^i is determined according to the manner in which the value of C^i compares with that of T^i. In other words, the effect of C^i on P^i is always diluted such that a change in the value of C^i by some constant e, will always result in the decrease of P^i by some constant $q < e$. Accordingly,

$$\left.\frac{\partial P^i}{\partial T^i}\right|_{C^i=0} = 1 \geq -\left.\frac{\partial P^i}{\partial C^i}\right|_{C^i=0} \tag{13.11}$$

for any $T^i \geq 0$. With the above properties in mind, the following proposition may be established.

Proposition 13.1. Let $P^i(T^i, C^i)$ be any function that satisfies the properties described in Assumption B1 with respect to any endeavour that satisfies Assumption B1–B2. Then the value of P^i will be more sensitive to changes in T^i than to changes in C^i, for all $T^i \geq 0$ and all $C^i > 0$.

Proof. According to (13.9), the partial derivative $\partial P^i/\partial T^i$ is both strictly positive and strictly non-decreasing as a function of (increasing) T^i (for any $C^i \geq 0$). Furthermore, (13.10) implies that the negated partial derivative $-\partial P^i/\partial C^i$ is both strictly positive and strictly non-increasing as a function of (increasing) C^i (for any $T^i > 0$). Therefore, if, for any fixed value $T^i \geq 0$, the inequality

$$\frac{\partial P^i}{\partial T^i} \geq -\frac{\partial P^i}{\partial C^i} \tag{13.12}$$

holds for some value $C^i = C^i_*(T^i)$, it follows from the properties in (13.9)–(13.10) that the inequality also holds for that fixed value T^i and all $C^i \geq C^i_*(T^i)$. It follows from (13.11), therefore, that the inequality (13.12), in fact, holds universally for all $T^i \geq 0$ and all $C^i > 0$. □

To confirm the significance of Proposition 13.1 it is necessary to illustrate that there are ample real-world examples of endeavours that satisfy Assumptions B1–B2, and additionally, that it is possible to employ Proposition 13.1 in the context of such systems in a manner that is of value. To address the former matter, consider the practical implications of Assumption B1. Essentially, the assumption admits any endeavour relating to a real-world system in which the entities contained therein are forced to coexist despite pursuing conflicting objectives. This kind of system may exist for several reasons, including, but not limited to: If multiple entities within a system are fundamentally linked to its purpose, if legislation governing such a system prohibits the exclusion of certain entities, or if no entity contained in such a system is capable of forcefully (and lawfully) extinguishing the presence of another.

It should be obvious that there are many real-world systems that exhibit the properties described above. Consider, for example, any system designed to deliver a product or a service to a set of customers. The entities contained in such a system can generally be partitioned into one of three categories, namely owners, employees and customers. It follows that the exclusion of any of the aforementioned role-players would force the related system into a state of disintegration — these entities are fundamentally linked to the purpose of the system in which they exist and must therefore coexist despite pursuing conflicting objectives. Another example of the kind of system described in Assumption B1 can be found in the fishing industry related to numerous coastal countries around the world. A reality commonly associated with these countries is the presence of coastal villages in which people survive by means of informal fishing. Unfortunately, more often than not, these villages come into conflict with commercial fishing firms operating along the same coastline [253]. According to international regulations, the amount of fishing allowed in a certain area is regulated by a strict quota system [252]. As such, commercial fishing firms that catch fish for profit, and small-scale fishermen who catch fish to survive are typically forced to share the available quota [253]. Obviously, either entity would be better off without the other, but because of the official regulations of the system in which they exist, they are forced to coexist despite pursuing conflicting objectives.

The final assumption, namely Assumption B2, has a rather one-dimensional effect on the kind of real-world system upon which Proposition 13.1 may be applied. It simply states that it must be possible to change the value of T^i without altering the value of C^i with respect to a certain objective i. Essentially, this implies that the pool of power with which entities may alter the measure of T^i can never be limited to the power contained in C^i, and *vice-versa*. There are very few real-world systems that fail to satisfy Assumption B2. More importantly, none of the systems presented to be descriptive of Assumption B1, fail to satisfy the properties stipulated by Assumption B2.

As stated, the final step in validating the significance of Proposition 13.1 is illustrating the utility thereof in the context of the real-world systems delineated above. Therein lies a considerable

difficulty — there are several technicalities, relevant to the interaction of competing entities, that are not addressed in the model derived above. To elucidate, consider that the difficulty of altering the value of T^i and that of C^i will likely be dissimilar in practice, but that the reality of such variation is not addressed in the model. In other words, although P^i has been proven to be more sensitive to changes in T^i than to changes in C^i, thereby suggesting that an entity should rather attempt to increase the value of T^i, than decrease the value of C^i in order to increase that of P^i, it may, in practice, be more effective to do just the opposite. Accordingly, until the mathematical model presented in this chapter is developed further to incorporate such matters, it is not possible to illustrate the practical applicability of Proposition 13.1 and, as a result, it is not possible to fully validate the significance thereof.

To address this shortcoming, the remainder of this chapter is dedicated to an extension of the model derived above by considering phenomena that participate in the act of altering the measure of power associated with the realisation of objectives pursued in the context of an endeavour. The first step taken to that end is the introduction of a slight modification to several elements presented thus far in the section that follows.

13.6 Extension of the model into the temporal domain

In order to capture (to some extent) the mechanisms involved in the alteration of the measure of power related to the pursuit of an objective in the context of an endeavour, a slight modification to several of the expressions introduced in the sections prior is required. More specifically, in order to model the notion of change, it is necessary to incorporate the notion of elapsed time — change necessitates, per definition, the reality of passing time. In other words, it is no longer sufficient to describe the measure of power and/or benefit associated with the pursuit of an objective without addressing the temporal dimension related to it. Table 13.1 contains an elucidation of the manner in which several parameters and variables introduced thus far are modified in order to be of utility in the remainder of this chapter.

TABLE 13.1: *A list of modified mathematical elements employed in the remainder of this chapter in order to model temporal changes to the dynamics of power in the context of an endeavour.*

Element	Description
P_x^{it}	The measure of power contributed toward the realisation of an objective $i \in \mathcal{I}_x$ by an entity x at time t
δ^{ijt}	The degree of correlation between two objectives $i, j \in \mathcal{I}$ at time t
T^{it}	The total measure of power working toward the realisation of an objective i at time t
C^{it}	The total measure of power working against the realisation of an objective i at time t
P^{it}	The measure of effective power associated with the pursuit of an objective i at time t
G^{it}	The measure of gain expected in the pursuit of an objective i if the power dynamics of an endeavour were to remain unchanged from time t onwards

Employing the modified elements presented above, the magnitude of power T^{it} working *toward* the realisation of an objective i at any time $t \in [0, s]$, where $t = 0$ represents the beginning and

$t = s$ the end of an endeavour, can now be represented by

$$T^{it} = \sum_{x=1}^{m} \sum_{j \in \mathcal{I}_x^+} \delta^{ijt} P_x^{jt}, \qquad i \in \mathcal{I}, \tag{13.13}$$

and the measure of power C^{it} working *against* the realisation of an objective i at any time $t \in [0, s]$ can now be represented by

$$C^{it} = \left| \sum_{x=1}^{m} \sum_{j \in \mathcal{I} \setminus \mathcal{I}_x^+} \delta^{ijt} P_x^{jt} \right|, \qquad i \in \mathcal{I}. \tag{13.14}$$

As stated in §13.5 (and obvious from the above expressions) an entity x can go about changing the measure of power working toward and/or against its objectives in one of two ways: By altering the degree of correlation δ^{ijt} between certain objectives $i, j \in \mathcal{I}$, or by altering the measure of power P_x^{jt} possessed by certain entities $x \in \{1, \ldots, m\}$, in such a way that T^{it} and C^{it} relative to the objectives in $\mathcal{I}_x$ increase and decrease, respectively.

The model extension presented in the section that follows addresses only the dynamics related to altering the magnitude of power possessed by certain entities in an endeavour, and that the reality of altering objective correlation is deemed work that may be pursued in future. The structure of the remainder of this chapter is thus as follows: First, an extended model aligned with the above specification is presented in §13.7, and, only thereafter, is an attempt made in §13.8 to generalise Proposition 13.1 taking into consideration the additional dynamics considered.

13.7 A mathematical description of the extended model

In order to facilitate a mathematical description of the mechanisms that govern the possession of power in the context of an endeavour, a strategy similar to the one adopted in the discussion of the original model in §13.2 and §13.3 is utilised — the complexity of matters considered is increased incrementally. Accordingly, the discussion commences with a markedly more general form of an expression for P_x^{jt}, the amount of power possessed by an entity x with respect to the pursuit of an objective $j \in \mathcal{I}_x^+$ at any time t during the course of an endeavour. In particular, it is assumed that

$$P_x^{jt} = \Psi_x^{jt} P_x^t \quad \text{for all} \quad t \in [0, s],$$

where $\Psi_x^{jt} \in [0, 1]$ represents the proportion of the power of entity x that is actually of utility in the pursuit of objective $j \in \mathcal{I}_x^+$, and P_x^t represents the total measure of the power of entity x irrespective of the pursuit of the objectives in $\mathcal{I}_x^+$. Although this expression is limited in its representation of the dynamics involved in the act of altering the measure of an entity's power, its adoption is of benefit as it facilitates the development of the elements contained therein. More specifically, the above expression may be refined by further developing the term P_x^t therein, so that

$$P_x^{jt} = \Psi_x^{jt} \left(P_x^0 + \int_0^t \Delta_x^\tau d\tau \right) \quad \text{for all} \quad t \in [0, s], \tag{13.15}$$

where P_x^0 represents the total measure of power possessed by an entity x at time $t = 0$ (*i.e.* the start of the endeavour), and $\Delta_x^\tau \in (-P_x^0, \infty)$ represents any alteration made to the amount of power possessed by entity x at time $\tau \in [0, t]$. Notably, this form of expression for P_x^{jt} reveals a term that is of particular interest (*i.e.* the term Δ_x^τ) as it presents the mechanism by which

the alteration of an entity's power may be induced. Put another way, the task of explicating Δ_x^τ and the task of mathematically describing the elements involved in the alteration of an entity's power are, for the most part, one and the same. Before embarking on an exposition of this term, there is a technicality concerning the constituent elements that should be noted.

In §9.2.1, the reader was briefly introduced to the notion of a power-play, then defined to constitute the very mechanism by which one entity may attempt to influence another. In this section, the topic of exerting influence is revisited with the introduction of a collection of dynamics related thereto, but the link between these dynamics, and the existence of power-plays has not yet been addressed. Accordingly, consider the similarity between the definition of the term Δ_x^τ and that of a power-play. The term Δ_x^τ has been defined to contain the alteration of the power of an entity x in the context of an endeavour, and the notion of a power-play as constituting the practical mechanism by which the alteration of the power of an entity x may take place. Evidently, by linking the underlying coherence of these definitions, Δ_x^τ can be viewed as representing the effect (in terms of power possessed) of any and all power-plays executed at time τ upon entity x. It thus follows that the term Δ_x^τ should be defined with respect to the dimensions that characterise power-plays. That is, their form, denoted by $y \in \{\text{coercive, suasive}\}$, and their source, denoted by $z \in \{\text{structural, customary, status, necessity, incapasitation}\}$[7].

13.7.1 The elements involved in the alteration of power

To facilitate the reader's understanding of the rather convoluted collection of elements that constitute the term Δ_x^τ, the utility of an analogous model, based on the laws of mechanics, is utilised. To avoid confusion, note that despite certain advantageous similarities, the model presented below is not considered a replication of the dynamics of an endeavour. The purpose thereof is purely to aid in general understanding by providing a manner in which to conceptualise the dynamics involved in the alteration of an entity's power[8]. With this in mind, the physical model to which the interaction of the elements contained in Δ_x^τ is considered analogous, is that of competing entities attempting to move two discrete, rectangular shaped objects on a non-smooth surface[9], in opposing directions, where the displacement of the one object (from some original position) represents an alteration in the measure of power possessed by one entity, say entity x, and the displacement of the other represents an alteration in the measure of power possessed by a second entity (in competition with the first)[10]. More formally, consider Figure 13.9 depicting a selection of the dynamics related to the movement of one of these objects (the one representing the power dynamics[11] of entity x, for argument sake). In the remainder of this section, the elements depicted in Figure 13.9, along with the dynamics they represent in the context of an endeavour, are introduced incrementally. To that end, consider the hypothetical scenario presented below.

Imagine that the aforementioned entity x (contained in some endeavour) does not possess the power it requires to attain its goals and, as such, undertakes to alter its position in the distribution of power related to the endeavour in question. Notably, there are two ways in which entity x may go about doing so: By increasing the measure of its own power, thereby increasing T^{it} for certain objectives $i \in \mathcal{I}_x$, or by decreasing the measure of power possessed by its opponent,

[7] For a review on the dimensions of a power-play see §9.2.1.

[8] Notably, there are several dynamics relevant to the analogous model that are not addressed as they simply lack utility in the explication of the phenomena of interest.

[9] Note that these hypothetical objects are sliding along a groove such that their orientation remains fixed.

[10] Each of the aforementioned entities may only attempt to move a single object at any one point in time.

[11] In other words, there exists a second rectangular object, in an identical setup to the one in Figure 13.9, that relates to the power dynamics of the competing entity, but that is not illustrated.

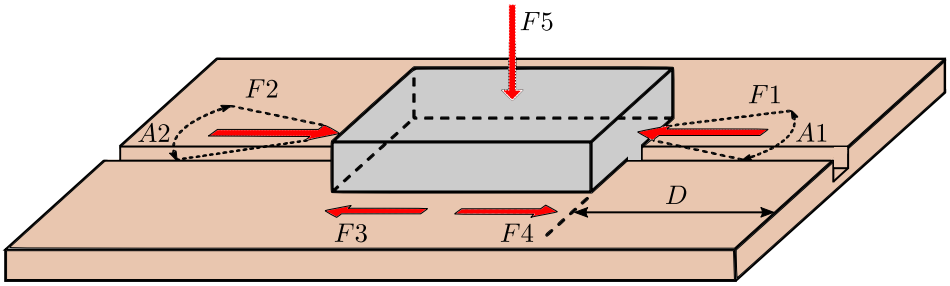


FIGURE 13.9: *A model, based on the laws of physics, analogous to the actions of entities attempting to alter the amount of power possessed by an entity x in the context of an endeavour.*

thereby decreasing C^{it} for certain objectives $i \in \mathcal{I}_x$. As the dynamics involved in these courses of action are equivalent, it is assumed (without loss of applicability to the alternative) that entity x selects to attempt to alter the measure of its own power and not that of its opponent[12] (*i.e.* induce a positive value of Δ_x^τ over some time interval). Notably, as the laws of mechanics dictate that altering the position of an object, like the one depicted in Figure 13.9, requires the exertion of a certain measure of energy, the mechanisms that govern an endeavour dictate that in order to aggregate more power, an entity, like the aforementioned entity x, must engage in some form of activity (*i.e.* the execution of a power-play) and thereby exert a measure of influence over some time interval toward achieving that goal.

It does not require any mental gymnastics to assert that the same reality holds for the entity opposing entity x. Notably, for the assumed case in which entity x attempts to increase its own power, such opposition entails one of two courses of action: Inhibiting the aggregation of the power of entity x, or following suite by attempting to aggregate its own power. Assuming that the first of these courses of action is pursued, it follows that the success thereof depends upon how the measure of influence exerted by entity x over a time period leading up to, and at time t, denoted by h_{xyz}^t, compares with the measure of influence exerted over the same time period, but by its competitor, say entity r, denoted by f_{ryz}^t. In the context of Figure 13.9, this dynamic is considered analogous to how the magnitude of the force $F1$ exerted to shift the depicted object to the left, compares with the magnitude of the force $F2$ exerted to shift the depicted object to the right[13] (at time t). Evidently, the interplay of the forces $F1$ and $F2$ plays a central role in determining whether the object in Figure 13.9 moves to the left, to the right, or not at all.

Despite its centrality to the successful alteration of the power dynamics in an endeavour, an unconstrained exertion of influence does not guarantee success in such an undertaking. Along with the measure of exerted influence, there are two additional, equally effectual phenomena that characterise the execution of any power-play. The first relates to an entity's skill in the execution thereof and, the second, to the degree of correlation between the power-play in question and the realisation of some predefined goal. These phenomena, which are both considered somewhat analogous to the angles $A1$ and $A2$ in Figure 13.9, are more formally described below, starting with the first.

[12] In the context of the model in Figure 13.9, this is analogous to entity x electing to dedicate its energy toward shifting the object depicted in the figure which, as stated, symbolises the power dynamics of entity x, as opposed to attempting to shift a second object (not depicted) symbolising the power dynamics of the competitor of entity x.

[13] Accordingly, the position of the object in the figure, denoted by D, is considered analogous to the amount of power possessed by entity x, the displacement of the object in the figure is considered analogous to the term Δ_x^τ, and the magnitude of the force $F1$ and the force $F2$ is considered analogous to the effort exerted by entities x and r in order to increase and decrease the measure of power possessed by entity x, respectively.

In essence, the relationship between skill and effectiveness is such that if an entity's skill in the execution of a power-play increases, the effectiveness of that power-play with respect to the realisation of its designed purpose, for an equal measure of exerted influence, must also increase, and similarly, if an entity's skill in the execution of such a manoeuvre decreases, so too must its effectiveness. In the context of the model in Figure 13.9, this dynamic is considered analogous to the effect of altering the angles $A1$ and $A2$ in Figure 13.9 such that they approach and recede from being equal to 90 degrees, respectively. It follows that if either the skill of entity x in the aggregation of power, denoted[14] by $k^t_{xyz} \in [0, \infty]$, or the skill of entity r in the degradation of power, denoted by $b^t_{rxy} \in [0, \infty]$, increases, so too must the effectiveness of the respective power-plays they execute toward the realisation of their respective goals.

Compared to the notion of skill, it is slightly more difficult to explicate the notion that alternative power-plays may exhibit differing levels of correlation with respect to the effect they have upon certain predefined outcomes. Toward elucidating this dynamic, note that the discussion thus far has, in fact, been governed by the assumption that whenever an entity attempts to alter the amount of power possessed by itself, or some other entity, the power-play it executes to that end is intended solely for that purpose. In practice, however, this is not always the case. It is possible, for example, that a power-play executed in an endeavour, only coincidently aligns somewhat to the act of altering the distribution of power contained therein, without that being its intended purpose. In such a case, it is likely that only a portion of the total measure of influence exerted to execute the power-play in question would actually contribute to the related alteration in power. The reality of this dynamic becomes even more apparent when one considers the fact that it is possible for some unknowing third party, not even a part of the endeavour in question, to alter the distribution of power contained therein unintentionally as a by-product of pursuing its own outcomes in a more general system (containing both the third party and the endeavour in question). In such a case, it is once again likely that only a portion of the influence exerted by the entity in question will actually contribute to the resulting alteration in power in the smaller endeavour considered — qualifying the aforementioned notion of correlation.

Like an entity's skill in the execution of a power-play, altering a power-play's correlation to the realisation of a certain outcome is considered analogous to altering the orientation of the forces $F1$ and $F2$ in Figure 13.9. Accordingly, the relationship between correlation and effectiveness is equivalent to that of skill and effectiveness. With respect to the aforementioned entities x and r, it thus follows that if either the correlation of a power-play executed by entity x to the aggregation of the power of entity x at time t, denoted by $\Omega^{xxt} \in [0, 1]$, or the correlation of a power-play executed by entity r with the degradation of the power of entity x at time t, denoted by $\Phi^{xrt} \in [-1, 0]$, increases, so too must the effectiveness of the power-plays they execute, respectively.

Denoted by the friction forces $F3$ and $F4$ in[15] Figure 13.9, there is a final dynamic to introduce before proceeding to the formal exposition of the term Δ^{τ}_{x}. Unlike any dynamic considered thus far, the effect thereof is considered analogous to the effect of friction, because, as friction is to movement, this dynamic is to the alteration of power in an endeavour. More specifically, the alteration of power in an endeavour does not take place in an unconstrained manner. Instead, the *environment*[16], both contained in and containing an endeavour, induces a regulatory constraint upon the alteration of power possessed by entities in an endeavour such that the alteration of one entity's power may be inherently more or less difficult than the alteration of another's. To illustrate the reality of this dynamic, recall the property that characterises an endeavour: The

[14] Recall that the subscripts y and z indicate the kind of power-play employed in the exertion of influence.

[15] Note that the gravitational force $F5$ in Figure 13.9 is included for the sake of accuracy and does not symbolise any of the dynamics relevant to the competitive interaction of entities in an endeavour.

[16] The environment comprises various mechanisms, including social norms, world-views and cultures.

coexistence of competing entities. Significantly, this property implies that there exists some form of influence in an endeavour (independent of the actions of a targeted entity) that is capable of constraining one entity's ability to decrease the power of another's. This form of restraint is considered a function of the amount of power a targeted entity possesses, the amount of influence an opposing entity exerts, and whether such influence is intended to increase or decrease the former entity's power. With respect to the aforementioned entities x and r, these dynamics imply that there exists a measure of environmentally induced influence that constrains[17] both the ability of entity x to aggregate power at time t, and the ability of entity r to decrease the power of entity x at time t, denoted by $\alpha_x^t \in [0, \infty)$.

13.7.2 Formally modelling the alteration of power

Now that all the dynamics involved in the alteration of the power dynamics of an endeavour have been introduced, it is possible to formally describe Δ_x^τ. For the sake of convenience, the collection of elements utilised in its description, are summarised in Table 13.2 below. The first step in describing Δ_x^τ is the introduction of the term W_{xyz}^τ which represents the net measure of influence experienced by an entity x as a result of any power-play executed in the form y and from the source z at time $\tau \in [0, t]$. Employing the elements in Table 13.2, this term may be expressed as

$$W_{xyz}^\tau = \sum_{n \in \mathcal{N}} h_{nyz}^\tau \Omega^{xn\tau} k_{nyz}^\tau + \sum_{r \in \mathcal{R}} f_{ryz}^\tau \Phi^{xr\tau} b_{ryz}^\tau \quad \text{for all} \quad \tau \in [0, t] \tag{13.16}$$

where $t \leq s$, the set $\mathcal{N}$ contains all entities n contributing to the aggregation of the power of an entity x, and the set $\mathcal{R}$ contains all entities r contributing to the degradation of the power of an entity x. Utilising (13.16), the term Δ_x^τ can be expressed as

$$\Delta_x^\tau = \sum_y \sum_z W_{xyz}^\tau - \alpha_x^\tau(W_x^\tau, P_x^{\tau-1}) \quad \text{for all} \quad \tau \in [0, t], \tag{13.17}$$

where α_x^τ is expressed as a function of W_x^τ and $P_x^{\tau-1}$, as the former term reveals the nature and severity of executed power-plays, and the latter indicates the measure of the power of the targeted entity x. Notably, as will be illustrated subsequently, the function $\alpha_x^\tau(W_x^\tau, P_x^{\tau-1})$ is design such that its sign convention is always opposite that of W_x^τ.

Now that an expression for the term Δ_x^τ has been established, it is finally possible to present the expression

$$P_x^{jt} = \Psi_x^{jt}\left(P_x^0 + \int_0^t \left(\sum_y \sum_z W_{xyz}^\tau - \alpha_x^\tau(W_x^\tau, P_x^{\tau-1})\right) d\tau\right), \tag{13.18}$$

for all $t \in [0, s]$, which represents (at least to some extent) the alteration of power of entity x during the course of an endeavour. As the reader may recall, $\Psi_x^{jt} \in [0, 1]$ represents the proportion of the power of entity x that is actually of utility in the pursuit of objective $j \in \mathcal{I}_x^+$, and P_x^0 represents the total measure of power possessed by entity x at time $t = 0$. Utilising this expression, a generalisation of Proposition 13.1 is attempted in the section that follows.

[17] In the context of the physical model in Figure 13.9, these phenomena can be understood in terms of the coefficients of friction of the non-smooth surface upon which the object in the figure moves, by stating that these coefficients of friction depend upon both the direction of movement, as well as the location (*i.e.* displacement D), of the object depicted.

TABLE 13.2: *A list of modified mathematical elements employed to model temporal changes in the dynamics of power in an endeavour.*

Element	Description
h^t_{nyz}	The measure of influence contributed to the aggregation of an entity's power by an entity n by executing a power-play of form y and from source z at time $t \in [0, s]$
f^t_{ryz}	The measure of influence contributed to the degradation of an entity's power by an entity r by executing a power-play of form y and from source z at time $t \in [0, s]$
k^t_{nyz}	The level of skill possessed by entity n in the execution of a power-play of form y and from source z at time $t \in (0, t]$
b^t_{ryz}	The level of skill possessed by an entity r in the execution of a power-play of form y and from source z at time $t \in [0, s]$
Ω^{xnt}	The correlation between a power-play executed by an entity n to the aggregation of the power of an entity x at time $t \in [0, s]$
Φ^{xrt}	The correlation between a power-play executed by an entity r to the aggregation of the power of an entity x at time $t \in [0, s]$
α^t_x	The environmentally induced resistance to the alteration of the power of an entity x at time $t \in [0, s]$

13.8 A generalisation of Proposition 13.1

At the conclusion of §13.5 it was evident that, in order to validate the practical applicability of Proposition 13.1, the model presented in §13.2 and §13.3 had to be extended. The primary motivation for the extension is that the original model limited the applicability of the aforementioned proposition to endeavours characterised by the rather uncommon property that equal influence exerted toward the modification of either the term T^{it} or the term C^{it} results in an equal change in their respective measures. Accordingly, the preceding section was dedicated to a model extension, capable of replicating (to some extent) the alteration of the distribution of power in an endeavour such that Proposition 13.1 may be generalised beyond the scope of the aforementioned property. The purpose of the present section is to utilise the extended model in an attempt to identify the respective states in an appropriate state space, with respect to which Proposition 13.1 may be generalised, thereby validating its usefulness[18].

Considering Proposition 13.1, it follows that the utility thereof may be generalised to any state for which one can illustrate that the effectiveness of influence exerted toward increasing one's power, is greater than, or at least equal to, the effectiveness of influence exerted toward decreasing the power of one's opponents (in terms of objective attainment). Toward describing this property mathematically, consider the hypothetical scenario below.

> During the course of an endeavour, two entities, entity a and entity b, find themselves competing with one another as a result of the objectives they pursue, characterised by set $\mathcal{I}_a \in \{A, B\}$ and the set $\mathcal{I}_b \in \{D\}$, respectively, where $\delta^{ABt} > 0$ and $\delta^{ADt} < 0$.

[18]The model extension presented includes only *entity-induced* alterations to the distribution of power in an endeavour. Power changes resulting from external forces such as natural disasters, are not included in the scope of this work.

Assuming that only the aforementioned entities a and b are able to alter the distribution of power among them (the effect of additional entities is addressed subsequently), it follows that the level of attainment, G^{At}, associated with the pursuit of objective A at time $t \in [0, s]$ is a function of $P^{At}(T^{At}, C^{At})$, the measure of power effective in the realisation thereof and, as such, depends upon how the measure of

$$T^{At} = (\delta^{AAt}\Psi_a^{At} + \delta^{ABt}\Psi_a^{Bt})\Bigg(P_a^0 + \int_0^t \Bigg(\sum_y \sum_z \Bigg(h_{ayz}^\tau \Omega^{aa\tau} k_{ayz}^\tau + f_{byz}^\tau \Phi^{ab\tau} b_{byz}^\tau\Bigg) - \alpha_a^\tau(W_a^\tau, P_a^{\tau-1})\Bigg) d\tau\Bigg), \quad (13.19)$$

for all $\tau \in [0, t]$ compares with that of

$$C^{At} = \delta^{ADt}\Psi_b^{Dt}\Bigg(P_b^0 + \int_0^t \Bigg(\sum_y \sum_z \Bigg(h_{byz}^\tau \Omega^{bb\tau} k_{byz}^\tau + f_{ayz}^\tau \Phi^{ba\tau} b_{ayz}^\tau\Bigg) - \alpha_b^\tau(W_b^\tau, P_b^{\tau-1})\Bigg) d\tau\Bigg) \quad (13.20)$$

for all $\tau \in [0, t]$. In view of the elements that constitute these expressions, consider the courses of action available to entity a toward altering the distribution of power among itself and entity b such that an increase the level of attainment G^{At} associated with objective A is affected. If entity a elects to improve its ability to compete by increasing its power, it will have to engage in some activity (*i.e.* the execution of a power-play of a certain form and source) such that the value of h_{ayz}^τ, in the first expression above, is larger than zero for some period of time $[0, t]$. The intention of such a manoeuvre would be to induce a corresponding $\Delta_a^\tau > 0$, thereby resulting in an increase in T^{At}. Conversely, if entity a elects to alter the distribution of power among itself and entity b by decreasing the power of entity b, it will have to engage in some activity (execute a power-play of the same, or alternative, form and source) such that the value of f_{ayz}^τ, in the second expression above, is larger than zero for some period of time $\tau \in [0, t]$. The intention this time would be to induce a corresponding $\Delta_b^\tau < 0$, thereby resulting in an increase in C^{At}. Accordingly, the property governing the generalisation of Proposition 13.1 to certain states in respect of the hypothetical endeavour presented above, may be rewritten in the form of the inequality[19]

$$\frac{\partial T^{At}}{\partial h_{ayz}^\tau} \geq \left|\frac{\partial C^{At}}{\partial f_{ayz}^\tau}\right|. \quad (13.21)$$

In view of the above expression, the reader may suspect, and correctly so, that a similar strategy to that adopted in the original exposition of Proposition 13.1 may be adopted to facilitate its generalisation. That is, utilising the behaviour of certain partial derivates to uncover the most effective way for entities to compete, depending upon the features that characterise them and their opponents. Before proceeding to such an exposition, a pair of assumptions (and a resulting modelling approach) in respect of the environment's constraining effect upon the alteration of power in an endeavour are introduced. The first of these assumptions is as follows:

A7. *The constrained alteration of power.* The less power an entity possesses, the more constrained the degradation of its power such that the measure thereof can never be nullified, and conversely, the more power an entity possesses, the less constrained the aggregation thereof.

[19] This inequality is presented under the assumption that entity A exerts influence in the form of h_{ayz}^τ and f_{ayz}^τ by utilising power-plays that are of the same form and source.

Although not a perfect representation of reality, it is deemed reasonable to assume that less powerful entities would find it more difficult to aggregate power (and *vice-versa*), and that the degradation of an entity's power becomes increasingly more difficult as the measure thereof approaches zero (this is, in fact, a property of an endeavour). The second assumption is:

A8. *The non-discriminant nature of the environment.* The constraining response of the environment to the attempted alteration of the power of an entity is considered invariant irrespective of the nature of the manoeuvre utilised to that end (*i.e.* whether its intended purpose is to increase or to decrease the power of an entity).

In view of the above assumptions, and the manner in which they limit the environment's response to the alteration of power, it follows that there are a considerable number of functions applicable to modelling the relationship between the environment and the alteration of power. Among the collection of such functions, there is one that is both particularly well-suited and rather uncomplicated, commonly called a sigmoid function. Using this kind of function, the environment's response α_x^t to an attempted alteration of the power of an entity x, is expressed as

$$\alpha_x^t = \mu(P_x^{t-1})W_x^t \quad \text{for } W_x^t \in (-\infty, \infty), \tag{13.22}$$

where

$$\mu(P_x^{t-1}) = \frac{1}{1 + e^{\eta P_x^{t-1} - \varkappa}},$$

where the parameter $\eta > 0$ alters the slope of the function μ, while the parameter $\varkappa > 0$ shifts the curvature thereof to the left or to the right, as depicted in Figure 13.10(a) and Figure 13.10(b), respectively. Notably, these elements are endeavour-specific and should be selected in accordance with the properties delineated in Assumption A7.

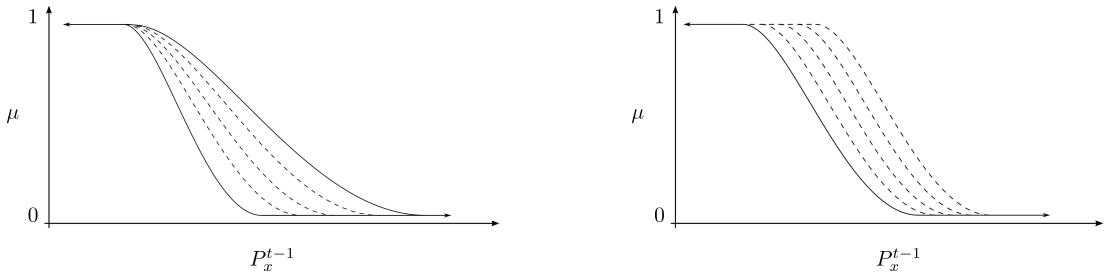


FIGURE 13.10: *(a) The effect upon $\mu(P_x^{t-1})$ when increasing η for a fixed value of $\varkappa$, and (b) the behaviour of $\mu(P_x^{t-1})$ when increasing $\varkappa$ for a fixed value of η.*

Now that the above technicalities have been made salient, the prerogative is to investigate the applicability of Proposition 13.1 to the state space corresponding to an endeavour. To that end, the expression in (13.22) is substituted for α_a^τ in (13.19) such that

$$\frac{\partial T^{At}}{\partial h_{ayz}^\tau} = (\delta^{AAt}\Psi_a^{At} + \delta^{ABt}\Psi_a^{Bt}) \int_0^t \left(\Omega^{aa\tau} k_{ayz}^\tau - \mu(P_a^{t-1})\Omega^{aa\tau} k_{ayz}^\tau \right) d\tau,$$

and similarly, the expression in (13.22) is substituted for α_b^τ in (13.20), such that

$$\frac{\partial C^{At}}{\partial f_{ayz}^\tau} = \delta^{ADt}\Psi_b^{Dt} \int_0^t \left(\Phi^{ba\tau} b_{ayz}^\tau - \mu(P_b^{t-1})\Phi^{ba\tau} b_{ayz}^\tau \right) d\tau.$$

Substituting the above expressions into (13.21), the inequality governing the generalisation of Proposition 13.1 may be expressed as

$$(\delta^{AAt}\Psi_a^{At} + \delta^{ABt}\Psi_a^{Bt}) \int_0^t \left(\Omega^{aa\tau} k_{ayz}^{\tau} - \mu(P_a^{t-1})\Omega^{aa\tau} k_{ayz}^{\tau}\right) d\tau \geq \left|\delta^{ADt}\Psi_b^{Dt} \int_0^t \left(\Phi^{ba\tau} b_{ayz}^{\tau} - \mu(P_b^{t-1})\Phi^{ba\tau} b_{ayz}^{\tau}\right) d\tau\right|.$$

Note that the term $\Phi^{bat} b_{ayz}^{\tau} - \mu(P_b^{t-1})\Phi^{ba\tau} b_{ayz}^{\tau}$ in the above inequality is, per definition, positive[20]. Consequently, the above expression may be modified such that the absolute value does not include this term. Doing so, the inequality may be rewritten as

$$\frac{\delta^{AAt}\Psi_a^{At} + \delta^{ABt}\Psi_a^{Bt}}{|\delta^{ADt}\Psi_b^{Dt}|} \geq \frac{\int_0^t (\Phi^{ba\tau} b_{ayz}^{\tau} - \mu(P_b^{t-1})\Phi^{ba\tau} b_{ayz}^{\tau}) d\tau}{\int_0^t (\Omega^{aa\tau} k_{ayz}^{\tau} - \mu(P_a^{t-1})\Omega^{aa\tau} k_{ayz}^{\tau}) d\tau}. \tag{13.23}$$

Although this form of the inequality does little to reveal any general insight concerning the generalisation of Proposition 13.1, it does facilitate a case-by-case enquiry into each state of the state space corresponding to an endeavour. To elucidate, by partitioning the state space in question according to the elements contained in the above inequality, a state specific enquiry may be conducted according to the features that characterise the resulting states. Accordingly, the state space corresponding to an endeavour is partitioned in the manner depicted in Table 13.3. As may be seen in the table, the degree to which objective A is correlated with the objectives in $\mathcal{I}_a \cup \mathcal{I}_b$ and the proportion of the power of the entities pursuing these objectives that is useful to their realisation, is adopted as the first feature on which to partition the state space. The state space is then partitioned according to the distribution of power among entities a and b, and then according to the ability of entity a to compete by utilising the courses of action available to it. Based on this partitioning, the applicability of Proposition 13.1 is investigated individually for the respective states in the state space of interest in the sections that follow.

13.8.1 State 1

In the context of State 1, characterised by the property that $\delta^{AAt}\Psi_a^{At} + \delta^{ABt}\Psi_a^{Bt} \geq |\delta^{ADt}\Psi_b^{Dt}|$, it follows that the inequality in (13.23) will hold if

$$\int_0^t \left(\Omega^{aa\tau} k_{ayz}^{\tau} - \mu(P_a^{t-1})\Omega^{aa\tau} k_{ayz}^{\tau}\right) d\tau \geq \int_0^t \left(\Phi^{ba\tau} b_{ayz}^{\tau} - \mu(P_b^{t-1})\Phi^{ba\tau} b_{ayz}^{\tau}\right) d\tau. \tag{13.24}$$

The question of interest is thus for which substates in State 1 this is indeed the case.

State 1.1

As depicted in Table 13.3, State 1.1 is characterised by the Property that $P_a^{t-1} = P_b^{t-1}$. This is a rather convenient property as it implies that, for an equal measure of influence exerted, the environment enacts an equal measure of restraint upon the alteration of the power of entity a and the power of entity b (*i.e.* $\mu(P_a^{t-1}) = \mu(P_b^{t-1})$). With this in mind, consider the first substate of State 1.1, State 1.1.1, characterised by the property that $\int_0^t (\Omega^{aa\tau} k_{ayz}^{\tau} d\tau = \int_0^t \Phi^{ba\tau} b_{ayz}^{\tau}) d\tau$. In combination with the aforementioned property in respect of the environment's restraint upon the

[20] The term $\mu(P_b^{t-1}) \in [0, 1]$.

TABLE 13.3: *A list of the states for which the generalisation of Proposition 13.1 is investigated.*

State	Property	State	Property	State	Property
1	$\delta^{AAt}\Psi_a^{At} + \delta^{ABt}\Psi_a^{Bt} \geq \lvert\delta^{ADt}\Psi_b^{Dt}\rvert$	1.1	$P_a^{t-1} = P_b^{t-1}$	1.1.1	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau = \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
				1.1.2	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau > \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
				1.1.3	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau < \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
		1.2	$P_a^{t-1} > P_b^{t-1}$	1.2.1	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau = \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
				1.2.2	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau > \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
				1.2.3	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau < \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
		1.3	$P_a^{t-1} < P_b^{t-1}$	1.3.1	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau = \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
				1.3.2	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau > \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
				1.3.3	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau < \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
2	$\delta^{AAt}\Psi_a^{At} + \delta^{ABt}\Psi_a^{Bt} < \lvert\delta^{ADt}\Psi_b^{Dt}\rvert$	2.1	$P_a^{t-1} = P_b^{t-1}$	2.1.1	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau = \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
				2.1.2	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau > \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
				2.1.3	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau < \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
		2.2	$P_a^{t-1} > P_b^{t-1}$	2.2.1	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau = \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
				2.2.2	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau > \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
				2.2.3	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau < \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
		2.3	$P_a^{t-1} < P_b^{t-1}$	2.3.1	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau = \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
				2.3.2	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau > \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$
				2.3.3	$\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau < \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$

alteration of power, it follows that under the conditions of State 1.1.1, the inequality in (13.24) reduces to

$$\int_0^t \left(\Omega^{aa\tau}k_{ayz}^{\tau} - \mu(P_a^{t-1})\Omega^{aa\tau}k_{ayz}^{\tau}\right)d\tau = \int_0^t \left(\Phi^{ba\tau}b_{ayz}^{\tau} - \mu(P_b^{t-1})\Phi^{ba\tau}b_{ayz}^{\tau}\right)d\tau,$$

thereby confirming that Proposition 13.1 may be generalised thereto. The inquiry into the remaining substates in State 1.1, State 1.1.2 and State 1.1.3, is facilitated by rewriting the inequality in (13.24) in the form

$$(1-\mu(P_a^{t-1}))\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau - (1-\mu(P_b^{t-1}))\int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau \geq 0.$$

Evidently, the fact that State 1.1.2 is characterised by the property that $\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau > \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$, and State 1.1.3 by the property that $\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau < \int_0^t \Phi^{ba\tau}b_{ayz}^{\tau}d\tau$, reveals that the above inequality will hold under the conditions of State 1.1.2, but not those of State 1.1.3.

State 1.2

If characterised by State 1.2, entity a is considered the more powerful entity, which, according to Assumption A7, implies that $\mu(P_a^{t-1}) < \mu(P_b^{t-1})$. In view of the exposition of State 1.1, the applicability of Proposition 13.1 to both State 1.2.1 and State 1.2.2 is thus apparent. Its applicability to State 1.2.3, however, is not. To facilitate the enquiry into State 1.2.3, the inequality in (13.24) is rewritten in the form

$$\frac{\int_0^t \Omega^{aa\tau}k_{ayz}^{\tau}d\tau}{\int_0^t \Phi^{bat}b_{ayz}^{\tau}d\tau} \geq \frac{1-\mu(P_b^{t-1})}{1-\mu(P_a^{t-1})}. \tag{13.25}$$

This form of the inequality reveals that although (13.24) does not hold for State 1.2.3 in general, it holds conditionally if the difference between the measure of power possessed by entity a and that possessed by entity b is sufficiently large, and/or the difference between the ability of entity a to increase its power and its ability to decrease the power of entity b is sufficiently small. Thus, Proposition 13.1 is deemed conditionally applicable to State 1.2.3.

State 1.3

Contrary to the aforementioned state, in the context of State 1.3 entity a is considered the less powerful entity. As a result, the environment's restraint upon the alteration of the power of Entity a is more severe than its restraint upon the alteration of the power of entity b (*i.e.* $\mu(P_a^{t-1}) > \mu(P_b^{t-1})$). Based upon the exposition in the case of State 1.1, it is apparent that Proposition 13.1 cannot be generalised to either State 1.3.1 or State 1.3.3. Its applicability to State 1.3.2, however, is conditional. To elucidate, consider the expression in (13.25). This form of the inequality makes it clear that under the conditions of State 1.3.2, (13.24) will hold if the difference between the amount of power possessed by entity a and that possessed by entity b is sufficiently small, and/or if the difference between the ability of entity a to increase its power and its ability to decrease that of entity b is sufficiently large. Consequently, Proposition 13.1 is deemed conditionally applicable to State 1.3.2.

13.8.2 State 2

State 2 characterises the decision instance facing entity a, if the objectives pursued by entity a are weakly correlated, few in number, and such that the power it possesses is of limited use to their realisation, and furthermore, if the objectives pursued by the opponent[21] entity a targets (entity b in this case) are considerable in number, strongly negatively correlated with objective A, and such that a significant portion of the opponent's power is of utility to their realisation.

State 2.1

In the context of State 2.1, entities a and b are considered equally powerful, which, as stated in §13.8.1, implies that the environment's restraint upon the alteration of the power of either these entities is equivalent (*i.e.* $\mu(P_a^{t-1}) = \mu(P_b^{t-1})$). That being the case, there is only one substate in State 2.1 for which the inequality in (13.23) may hold[22], namely State 2.1.2. This state is characterised by the property that the ability of entity a to increase its power is greater than its ability to decrease the power of entity b. Inequality (13.23) may be rewritten as

$$\frac{\delta^{AAt}\Psi_a^{At} + \delta^{ABt}\Psi_a^{Bt}}{|\delta^{ADt}\Psi_b^{Dt}|} \geq \frac{(1-\mu(P_b^{t-1}))\int_0^t \Phi^{ba\tau} b_{ayz}^{\tau} d\tau}{(1-\mu(P_a^{t-1}))\int_0^t \Omega^{aa\tau} k_{ayz}^{\tau} d\tau}, \tag{13.26}$$

which makes it clear that (13.23) will hold if the difference between the absolute value of the term $\delta^{ADt}\Psi_b^{Dt}$ and the term $\delta^{AAt}\Psi_a^{At} + \delta^{ABt}\Psi_a^{Bt}$ is sufficiently small, and/or if the difference between the ability of entity a to increase its power and its ability to decrease the power of entity b is sufficiently large. Accordingly, Proposition 13.1 is deemed conditionally applicable to State 2.1.2.

[21] That is, if the entity in question elects to alter the power dynamics of the endeavour of which it is a part by decreasing the power of an opponent.

[22] Note that in the context of State 2, the conditions stipulated in (13.24) are not sufficient to establish the applicability of Proposition 13.1.

State 2.2

When characterised by the conditions of State 2.2, entity a is considered the more powerful entity. This implies that $\mu(P_a^{t-1}) < \mu(P_b^{t-1})$ and, as such, it follows from the above discussion, concerning State 2.1, that Proposition 13.1 conditionally applies to all three substates of State 2.2. Consider first State 2.2.1, characterised by the property that the ability of entity a to increase its power equals its ability to decrease the power of entity b. Under these conditions, it follows from the expression in (13.26) that the inequality in (13.23) will hold if the difference between the absolute value of the term $\delta^{ADt}\Psi_b^{Dt}$ and the term $\delta^{AAt}\Psi_a^{At} + \delta^{ABt}\Psi_a^{Bt}$ is sufficiently small, and/or if the difference between the measure of power possessed by entity a and that possessed by entity b is sufficiently large. Concerning the remaining substates of State 2.2, it follows from the properties that characterise them that, in addition to the conditions stated in respect of State 2.2.1, the inequality in (13.23) will hold for State 2.2.2 if the difference between the ability of entity a to increase its power and its ability to decrease the power of entity b is sufficiently large, and for State 2.2.3 if the difference between these measures is sufficiently small.

State 2.3

Due to the properties that characterise State 2.3, Proposition 13.1 fails to apply to all but one of the substates contained therein. That is, State 2.3.2, characterised by the property the entity a is more able to increase the measure of its own power than it is to decrease that of entity b. According to this property, it follows by the expression in (13.26) that the inequality in (13.23) will hold under the conditions of State 2.3.2, if the properties characterising the interaction of entities a and b are such that the difference between the absolute value of the term $\delta^{ADt}\Psi_b^{Dt}$ and the term $\delta^{AAt}\Psi_a^{At} + \delta^{ABt}\Psi_a^{Bt}$ is sufficiently small, and/or if the difference between the measure of power possessed by entity a and that possessed by entity b is sufficiently small, and/or if the difference between the ability of entity a to increase its power and its ability to decrease the power of entity b is sufficiently large. Thus, Proposition 13.1 is deemed conditionally applicable to State 2.2.3.

13.8.3 The effect of additional entities

Before presenting a summary of the results pertaining to the above generalisation of Proposition 13.1, the effect of additional players upon the validity of the analysis should be addressed. Notably, the generalisation of Proposition 13.1 was conducted by considering the situation in which only two entities were considered affected by the alternative courses of action towards altering the distribution of power in an endeavour (*i.e.* entities a and b). It should be intuitively apparent, however, that the dynamics related to these courses of action will be subject to change if additional entities are taken in account. Consider, for example, the effect of including an entity c which pursues an objective E, where $\delta^{AEt} < 0$, and an entity d that pursues an objective F, where $\delta^{AFt} > 0$. In the presence of these entities,

$$\frac{\partial T^{At}}{\partial h_{ayz}^{\tau}} = (\delta^{AAt}\Psi_a^{At} + \delta^{ABt}\Psi_a^{Bt}) \int_0^t \left(\Omega^{aa\tau} k_{ayz}^{\tau} - \mu(P_a^{t-1})\Omega^{aa\tau} k_{ayz}^{\tau}\right) d\tau + \delta^{AEt}\Psi_d^{Et} \int_0^t \left(\Omega^{da\tau} k_{ayz}^{\tau} - \mu(P_d^{t-1})\Omega^{da\tau} k_{ayz}^{\tau}\right) d\tau$$

and

$$\frac{\partial C^{At}}{\partial f^{\tau}_{ayz}} = \delta^{ADt}\Psi^{Dt}_{b}\int_{0}^{t}\left(\Phi^{ba\tau}b^{\tau}_{ayz} - \mu(P^{t-1}_{b})\Phi^{ba\tau}b^{\tau}_{ayz}\right)d\tau + \delta^{AEt}\Psi^{Et}_{c}\int_{0}^{t}\left(\Phi^{ca\tau}b^{\tau}_{ayz} - \mu(P^{t-1}_{c})\Phi^{ca\tau}b^{\tau}_{ayz}\right)d\tau$$

for all $\tau \in [0, t]$. Each of the above expressions contain several parameters and variables related to either entity c or entity d, indicating that the generalisation of Proposition 13.1 depends upon the specific configuration of entities involved in an attempted alteration of power. Although this is somewhat inconvenient, conducting an analysis during which the applicability of Proposition 13.1 is determined for a specific decision instance (like the one above) is uninvolved. The generalisation of Proposition 13.1 to such instances is deemed work that may be pursued in future.

To conclude the discussion on the generalisation of Proposition 13.1 concerning the state space corresponding to the hypothetical endeavour containing entities a and b, the results related to this enquiry are summarised in Table 13.4. The following colour scheme is utilised: A state is coloured *green* if Proposition 13.1 may be generalised thereto, *yellow* if Proposition 13.1 is conditionally applicable thereto, or *red* otherwise.

TABLE 13.4: *A summary of the results pertaining to the generalisation of Proposition 13.1 in respect of the state space of an hypothetical endeavour containing two entities, a and b.*

State	Property	State	Property	State	Property
1	$\delta^{AAt}\Psi^{At}_{a} + \delta^{ABt}\Psi^{Bt}_{a} \geq \lvert\delta^{ADt}\Psi^{Dt}_{b}\rvert$	1.1	$P^{t-1}_{a} = P^{t-1}_{b}$	1.1.1	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau = \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
				1.1.2	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau > \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
				1.1.3	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau < \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
		1.2	$P^{t-1}_{a} > P^{t-1}_{b}$	1.2.1	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau = \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
				1.2.2	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau > \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
				1.2.3	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau < \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
		1.3	$P^{t-1}_{a} < P^{t-1}_{b}$	1.3.1	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau = \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
				1.3.2	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau > \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
				1.3.3	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau < \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
2	$\delta^{AAt}\Psi^{At}_{a} + \delta^{ABt}\Psi^{Bt}_{a} < \lvert\delta^{ADt}\Psi^{Dt}_{b}\rvert$	2.1	$P^{t-1}_{a} = P^{t-1}_{b}$	2.1.1	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau = \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
				2.1.2	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau > \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
				2.1.3	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau < \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
		2.2	$P^{t-1}_{a} > P^{t-1}_{b}$	2.2.1	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau = \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
				2.2.2	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau > \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
				2.2.3	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau < \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
		2.3	$P^{t-1}_{a} < P^{t-1}_{b}$	2.3.1	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau = \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
				2.3.2	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau > \int_\tau^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$
				2.3.3	$\int_0^t \Omega^{aa\tau}k^{\tau}_{ayz}d\tau < \int_0^t \Phi^{ba\tau}b^{\tau}_{ayz}d\tau$

13.9 Chapter summary

The mysteries surrounding the dynamics that govern the mechanisms of human interaction have often dominated the pursuit of scientific enquiry throughout history. The fact that this pursuit remains indefinite suggests that either its methods are failing, or that the phenomenon being studied is intractably complex. In view of the delineated history of this pursuit, the former

reason for its continuation is rejected. The work that has been presented in this chapter is the product of an attempt at moving slightly closer to an adequate description of the mysteries surrounding human interaction.

In view of the collection of assumptions made during the course of the development of the model presented in this chapter, the context upon which it is based can only be described as highly idealised. In accordance, any insight gained should be considered with caution.

The introduction of a governing assumption underlying the nature of human decision making in §13.1 was the first step of this attempt. This was followed by a mathematical description of the relationship between the mechanisms of expected gain and that of power possessed in the context of an endeavour in §13.2–13.3. Thereafter, a brief description of the effect of infeasibility upon this relationship was presented in §13.4. In §13.5, an analysis of the model was presented, and this was followed by an extension of the model in §13.6 and §13.7. Finally, §13.8 was dedicated to the generalisation of a result related to the manner in which entities compete in an endeavour.

Part III

Conclusion

CHAPTER 14

Dissertation summary and appraisal

Contents

This chapter consists of two sections. A summary of the research documented in this dissertation is provided in §14.1, and this is followed in §14.2 by an appraisal of contributions made in the dissertation.

14.1 Dissertation summary

Apart from the introductory chapter and its successor, the latter of which was dedicated to a discussion of the research methodology adopted, the work presented in this dissertation was organised into three parts, collectively comprising a further thirteen chapters. The first part, Part II, contained six chapters that were dedicated to a review of the academic literature relevant to the topic of this dissertation, in fulfilment of Objective I of §1.3. Part I opened in Chapter 3 with a brief introduction to the mechanisms of human cognition in §3.1, and this was followed by a discussion on the heuristics that are commonly employed in judgements related to uncertain events in §3.2. These heuristics were classified into three categories, namely simple heuristics, complex heuristics and hyper-heuristics. The aforementioned investigation set the stage for an account of the cognitive biases that render judgement heuristics imperfect in §3.4–3.8, after which Chapter 3 came to a close with a short summary in §3.9.

The next chapter, Chapter 4, was dedicated to an enquiry into the social nature of human decision making. After this topic was briefly introduced in §4.1 with a discussion on the paradoxical nature of group behaviour, the social dynamics related to decision making in the presence of relevant others was explicated in §4.2–§4.4. A central theme in the chapter was that group interaction and conflict go hand in hand — conflict between the needs of an individual, but more notably, the conflict that exists among interacting groups. Before the chapter concluded in §4.5 with a brief summary, it was illustrated that one of the driving forces behind this conflict is the naturally human tendency to protect what is experienced as an extension of the self, namely, the ingroup.

In accordance with the objectives pursued in the dissertation, Chapter 5 was dedicated to an enquiry into the field of debiasing. The chapter opened with a brief introduction to the field in §5.1, after which the prevailing schools of thought that underlie the methods employed therein

were reviewed in §5.2–§5.4. After a collection of methods based upon the reviewed schools of thought were reviewed in §5.5, the chapter came to a close with a short summary in §5.6.

Chapter 6 embodied quite a shift in the focus of discussion, opening in §6.1 with an account of the origins of systems theory. Excluding a short summary of matters discussed in §6.10, the chapter was essentially dedicated to an analysis of the paradigms that dominate systems theory, with the respective solution methodologies related thereto reviewed in §6.1–§6.9.

Chapter 7 opened with a brief introduction to the notion of dissimilar problem types in §7.1, and this was followed in §7.2–§7.4 by an evaluation of three unique problem typologies for the classification of real-world problems in terms of their effectiveness in eliciting the characteristics that are relevant to the process of real-world problem solving. The chapter finally came to a close with a short summary in §7.5.

Chapter 8, with which Part I concluded, was dedicated to a review of the methods that dominate the field of stakeholder selection theory. The chapter opened in §8.1 with a review of several methods that facilitate the identification of stakeholders, and this was followed in §8.2–§8.5 by an analysis of a collection of methods considered to facilitate the classification of stakeholders, and in §8.6 by a discussion of a collection of methods that facilitates an investigation into the network of relationships among stakeholders. In §8.7, a high-level comparison of the methods discussed was presented, after which the chapter was brought to a close with a brief summary in §8.8.

Part II of the dissertation consisted of a further five chapters, which started in Chapter 9 with a delineation of the meta-methodology proposed in fulfilment of Objectives II–VII of §1.3. The chapter opened with a brief overview of the phases that constitute the proposed methodology in §9.1, which set the stage for an in-depth discussion on each of these phases in the sections that followed. More specifically, §9.2 was dedicated to the introduction of a method for stakeholder selection, as well as an analysis of the sociological paradigms upon which the structuring methods of the systems literature are based, while §9.3 was reserved for the unveiling of a decision-making checklist designed to mitigate the irrational influence in human decision making. The chapter finally concluded with a concise summary of matters discussed in §9.4.

Chapter 10 contained an account of the first stage of evaluation of the proposed meta-methodology, involving its hypothetical application in the context of a theoretical case study in retrospect. The chapter commenced with an account of the selected context of application in §10.1, and this was then followed by the hypothetical application of the proposed meta-methodology in that particular context in §10.2. Thereafter, an additional phase of the meta-methodology, designed to address a potentially detrimental phenomenon uncovered during its hypothetical application, was introduced in §10.3. The chapter finally came to a close with a short summary in §10.4.

In Chapter 11 the proposed meta-methodology was subjected to the critique of an expert analyst. The chapter opened with a brief account of the feedback received from the analyst, in §11.1, and this was followed by the presentation of an updated version of the meta-methodology in §11.2. The chapter then came to a close with a short summary, in §11.3.

Chapter 12 was dedicated to the documentation of the final phase of evaluation of the proposed meta-methodology — its application to a practical case study. Accordingly, the chapter opened with a brief description of the context in which the meta-methodology was applied in §12.1, and this was followed with an account of its actual application in §12.2 and §12.3. The chapter came to a close with a brief summary, in §12.4.

In the final chapter of Part II, the subject matter became slightly more mathematical as Chapter 13 comprised an exposition of an analytical model developed for describing (to a certain

extent) the competitive interaction among the entities relevant to the context of application for which the proposed meta-methodology was designed, in fulfilment of Objective VIII of §1.3. The derivation of the model commenced with the introduction of rationality as a governing assumption in §13.1, and this was followed by a mathematical description of the relationship between the notions of power and potential gain in §13.2–§13.4. In §13.5, an analysis of the model was conducted and, as a result, the section concluded with an interesting result pertaining to the manner in which entities are to compete in the context of an endeavour (captured in Proposition 13.1). Next, an extension to the model was presented in §13.6 and §13.7, and this was followed in §13.8 by the generalisation of the aforementioned proposition to the state space induced by the extension. The chapter finally came to a close with a short summary in §13.9.

14.2 Appraisal of dissertation contributions

The main contributions of this dissertation were seven-fold. This section is dedicated to a documentation and an appraisal of these contributions.

Contribution 14.1 *A thorough investigation of the paradigms of enquiry relevant to the systems movement, as well as the solution methodologies contained within those paradigms, as documented in the academic literature.*

The systems movement embodies such a conglomeration of academic disciplines and fields of practice that it is almost impossible to make sense of the burgeoning expanse of literature related thereto. In order, therefore, to properly realise the pursuit of Objective V, it was first necessary to attempt to "bring order" to that literature. To that end, a delineation of the paradigms of enquiry that embody the systems movement, the methodological approaches related to each of those paradigms, and the methods that embody each of the approaches may be found in Chapter 6.

Contribution 14.2 *A thorough analysis of the dynamics relevant to the process of human decision making in the face of uncertainty, and in the context of social stimuli, as documented in the academic literature.*

Although real-world problem solving is fundamentally a human-centric endeavour, the dynamics that accompany that very quality are barely ever addressed in the structuring methodologies of systems theory and, as such, the mechanisms of human decision making were a central theme throughout the exposition of this dissertation. More specifically, the cognitive and social dynamics that constitute the mechanisms of human decision making were brought together in a unique and hopefully insightful manner in Chapters 3 and 4, respectively.

Contribution 14.3 *The development, and application, of a generic method for the identification of stakeholders that are relevant to a certain instance of real-world problem solving.*

Following an investigation of the relevant academic literature, it was discovered that solution methodologies dedicated to the formulation of real-world problems are typically void of guidance with respect to the identification of stakeholders. The neglect of this task was indeed a surprise as the appropriate selection of stakeholders depends first on their identification. In order, therefore, to fill this void in the methods of systems theory, a generic method for the identification of stakeholders, based upon the general criteria of power possessed and effect experienced, was proposed in §9.2.1. It was necessary to redefine the notions of power and effect in

order to construct this method, so as to elicit the mechanisms by which the experience of effect propagate, and finally, to provide an account of the criteria according to which the severity of effect experienced and the importance of power possessed may be determined. The efficacy of these contributions was first illustrated by means of a theoretical case study in respect of the fees must fall movement and, thereafter, by means of a practical case study conducted in the context of the South African energy sector.

Contribution 14.4 *A delineation of the effect of dissimilar states in which real-world systems exist upon the resolution of problems therein, as well as the development of a method designed to minimise the cost associated with one such state.*

Problem solution methodologies dedicated to the formulation of real-world problems unjustifiably neglect to address not only that real-world systems may exist in one of several states, but furthermore, that the state in which a system exists has a profound impact on the resolution of problems therein. Accordingly, to rectify this neglect, the meta-methodology proposed in this dissertation was dedicated, in part, to a portrayal of the states in which a system can exist. More specifically, the dissimilar states in which real-world systems exist, and the mechanisms by which these states come about, were first described, and then practically illustrated in §10.3. Notably, the undertaking to aid practitioners in dealing with the states in which real-world systems exist did not end with the aforementioned contribution. Based upon the body of knowledge uncovered, a method was designed and formally proposed for minimising the cost associated with one such state (*i.e.* a state of chaos).

Contribution 14.5 *An analysis of the sociological paradigms upon which the design of the structuring methods of the systems literature are based.*

To counteract the cumbersome manner in which real-world problem solvers tend to base their selection of solution methodologies upon their history of method use, and not according to how well the design of a particular method compliments the complexity of a certain task, an analysis of the sociological paradigms upon which the methods of systems theory are based, was presented in §9.2.3. This was, furthermore, accompanied by a description of the paradigm-based practical pitfalls associated with a number of structuring methodologies related to the systems movement in §9.2.4.

Contribution 14.6 *A decision-making checklist designed to mitigate the irrational influence in human decision making.*

All too often it is not the complex nature of problems, but the subtle dynamics that exist among and within problem solvers that result in the persistence of problems considered. The mitigation of such dynamics has, furthermore, proven remarkably difficult throughout the history of human decision making as they often operate in the subconscious aspect of the human mind. Accordingly, to address this aspect of human nature, it was necessary to adopt a strategy that did not involve the actual identification of the aforementioned dynamics, but simply the mitigation of their effect. The decision-making checklist presented in §9.3, constitutes the product of adopting this strategy, and comprises a framework designed to address the social and cognitive dynamics determined to be most adverse to the process of human decision making.

Contribution 14.7 *The development of a model describing mathematically the competitive interaction of entities relevant to the context for which the proposed meta-methodology was designed.*

The reasons for which the mathematical model presented in Chapter 13 is considered a contribution is twofold. For one, it resulted in the formalisation of Proposition 13.1, which embodies a useful step toward providing the competitive entities of real-world problem solving with case-specific guidance in respect of how to compete during an endeavour. Furthermore, the mathematical description of the competitive interaction of stakeholders contained in the model certainly goes some way towards unveiling the mechanisms that govern similar forms of interaction in general.

CHAPTER 15

Future work

Contents

During the process of conducting the research reported in this dissertation, several avenues of further investigation were noted and, as such, this chapter is dedicated to their delineation.

15.1 Suggestions related to scoping delineations

Two kinds of future work suggestions are contained in this chapter, and this section is dedicated to the first of these. More specifically, two avenues of further investigation related to the scope of the research carried out in the preceding chapters are discussed here. In each case, the suggestion is stated formally and then motivated briefly.

Suggestion 10.1 *Test the effectiveness of the method proposed to facilitate the selection of an appropriate response to a chaos-inducing entity by means of practically applying it in the context of a real-world endeavour.*

The meta-methodology proposed in §10.3 was developed in response to the recognised difficulty of responding effectively to an entity that utilises chaos as a means of influence in the context of an endeavour. Although the logic underlying the methodology was illustrated in the context of a theoretical case study and its working subjected to the scrutiny of an expert analyst, in order to further validate it, the methodology should be applied in the context of a real-world endeavour. Its evaluation by means of practical application thus forms part of the proposed future work.

Suggestion 10.2 *Broaden the scope of the proposed meta-methodology to include the process of response formulation by means of quantitative modelling.*

The meta-methodology proposed in this dissertation is built upon the synthesis of dissimilar system paradigms (inspired by the critical system school of thought). Accordingly, the proposed meta-methodology has been developed such that the scope of its application may be extended to include the process of response formulation, and thus the methods of hard systems thinking. To

elucidate, a natural extension to the work conducted in this dissertation would be to execute an analysis of the methods in the systems literature dedicated to the process of response formulation, with the aim of identifying the limitation thereof so that the proposed meta-methodology may be supplemented with the capacity to facilitate their use. Note that such an extension of the meta-methodology would also involve investigating and addressing the complexities associated with the process of transitioning from the problem formulation to the problem response phase in the context of real-world problem solving.

15.2 Natural extensions of the work presented

In this final section, six possible avenues of further investigation are suggested that follow as natural extensions to the work documented in this dissertation. In each case, the suggestion is again stated formally and then motivated briefly.

Suggestion 10.3 *Further develop the mathematical model presented in Chapter 13 to incorporate the capability of altering the amount of power associated with the pursuit of an objective by altering the degree of correlation between the objectives in $\mathcal{I}$.*

It seems reasonable to suggest that the degree of correlation between the objectives pursued in an endeavour may deviate from some baseline value during the course of its progression. Consider, for example, two entities attempting to redefine the outcomes they pursue such that the degree of correlation between the objectives that represent these outcome increases (or decreases) for their mutual benefit. With this kind of scenario in mind, consider the following question of interest: When is it more effective for an entity x to attempt to increase the degree of correlation between the objectives $\mathcal{I}_x$ it pursues and the objectives $\mathcal{I}_x^+$ positively correlated thereto, compared to decreasing the degree of correlation between the objectives in $\mathcal{I}_x$ and the objectives in $\mathcal{I}\backslash\mathcal{I}_x^+$ that are non-positively correlated thereto. Toward answering this question, the mathematical model presented in Chapter 13 may be generalised to incorporate the phenomena involved in the task of varying the degree of correlation among objectives such that the validity of Proposition 13.1 may be determined in respect of the state space induced by such an extension.

Suggestion 10.4 *Review the applicability of Proposition 13.1 to the state space induced by the assumption that the environment is biased against either the aggregation or the degradation of an entity's power.*

Proposition 13.1 was established under the assumption that the environment constrains the alteration of an entity's power in a manner that disregards the nature of the manoeuvre utilised to that end (*i.e.* whether its intended purpose is to increase or to decrease the power of the entity in question). Certain real-world systems may, however, be characterised by the property that this form of constraint depends upon the manoeuvre utilised. Consequently, investigating the applicability of Proposition 13.1 in the case where a unique function $\mu(P_x^{t-1})$ is utilised to determine the degree of constraint imposed upon influence exerted toward increasing and toward decreasing the power of an entity, is considered work that may pursued in future. Such an enquiry may provide insight into when it could be more effective for an entity to attempt to increase the measure of its own power, compared to decreasing the measure of an opponent's power, even when doing so is considered inherently more difficult.

Suggestion 10.5 *Extend Proposition 13.1 to instances of competition involving more than two entities.*

As illustrated in §13.8, the courses of action considered of utility to an entity in the pursuit of increasing its ability to compete in a particular instance of competition depends upon the number of entities involved. At present, the applicability of Proposition 13.1 is limited to such instances involving only two entities, but as endeavours are seldom limited only to two participants, it is probable that numerous entities may be involved in a particular instance of competitive interaction. Consequently, the generalisation of Proposition 13.1 to such instances is considered work that may be pursued in future. To that end, an exemplar of how such an analysis may be conducted is provided in §13.8.

Suggestion 10.6 *Further develop the decision-making checklist by considering additional dynamics, elaborating on the manner in which dynamics considered may practically manifest themselves, and enhancing the enquiry designed to mitigate their influence.*

The decision-making checklist presented in §9.3 consists of three components: A list of dynamics capable of inhibiting successful real-world problem solving, a description of the manner in which such dynamics may manifest themselves in a practical context, and finally, an enquiry designed to mitigate the unwanted influence associated with the dynamics considered. Accordingly, each of these components may further be developed. The decision making-checklist may be supplemented to incorporate the influence of dynamics not currently considered in order to address additional forms in which the dynamics currently considered could manifest themselves, and finally, to better facilitate the enquiry designed to mitigate the influence of such dynamics.

Suggestion 10.7 *Put together a catalogue of features commonly associated with real-world problems.*

As described in §9.2.1, establishing an exhaustive classification of the dimensions that underlie the human experience of effect is a task that has remained a persistent challenge in the academic literature. In like manner, in this dissertation, the mechanisms of effect were found to be so intricately entwined that it was not possible to untangle them in a detailed manner that maintained general applicability. In order to better aid practitioners in the development of a functional network related to a problem considered there is, however, an alternative approach that may be pursued. That is, the development of a catalogue of potential features from which practitioners may draw inspiration when attempting to identify features that are central to a problem considered. In this way, although the dimensions underlying the experience of effect remain unidentified, the features responsible for inducing such an effect may be made more salient.

Suggestion 10.8 *Further develop the methodological analysis of §9.2.3 by investigating the actual working of the structuring methods considered.*

In its current form, the methodological analysis conducted in §9.2.2 involves the consideration of only the sociological paradigms upon which the structuring methodologies of the systems literature are based, and not the actual working of the structuring methods themselves. Accordingly, the current methodological analysis may further be developed to include an exposition of the working of the methods considered, and the effect thereof upon the practical contexts in which they may be of utility.

References

[1] Abelson RP & Levi A, 1985, *Decision making and decision theory*, pp. 231–309 in Lindzey G & Aronson E (Eds), *Handbook of social psychology*, Random House, New York (NY).

[2] Ackermann F & Eden C, 2001, *Stakeholders matter: Techniques for their identification and management*, University of Strathclyde, Glasgow.

[3] Ackermann F & Eden C, 2010, *Strategic options development and analysis*, pp. 135–190 in Reynolds M & Holwell S (Eds), *Systems approaches to managing change: A practical guide*, Springer, London.

[4] Ackermann F & Eden C, 2011, *Strategic management of stakeholders: Theory and practice*, Long Range Planning, **44(3)**, pp. 179–196.

[5] Ackoff RL, 1962, *Scientific method: Optimizing applied research decisions*, Wiley, New York (NY).

[6] Ackoff RL, 1979, *The future of operational research is past*, Journal of the Operational Research Society, **30(2)**, pp. 93–104.

[7] Ackoff RL, 1981, *The art and science of mess management*, Interfaces, **11(1)**, pp. 20–26.

[8] Ackoff RL, 1988, *A theory of practice in the social systems sciences*, Systems Research and Behavioral Science, **5(3)**, pp. 241–246.

[9] African News Agency, 2016, *Calm down, fees aren't rising yet — Blade*, [Online], [Cited October 2018], Available from https://citizen.co.za/news/south-africa/1250834/calm-down-fees-arent-rising-yet-blade/.

[10] Ahlstrom D & Wang LC, 2009, *Groupthink and France's defeat in the 1940 campaign*, Journal of Management History, **15(2)**, pp. 159–177.

[11] Ajzen I, 1977, *Intuitive theories of events and the effects of base-rate information on prediction*, Journal of Personality and Social Psychology, **35(5)**, pp. 303.

[12] Albarracin D & Wyer Jr RS, 2000, *The cognitive impact of past behavior: Influences on beliefs, attitudes, and future behavioral decisions*, Journal of Personality and Social Psychology, **79(1)**, pp. 5–22.

[13] Albers MJ, 2002, *Design considerations for complex problem-solving*, Proceedings of the Society for Technical Communication Annual Conference, Rosemont (IL), pp. 217–222.

[14] Alderfer CP, 1983, *An intergroup perspective on group dynamics*, (Unpublished) Technical Report, New Haven Connecticut School of Organisation and Management, Yale University, New Haven (CT).

[15] Allaire JC & Marsiske M, 2002, *Well- and ill-defined measures of everyday cognition: Relationship to older adults' intellectual ability and functional status*, Psychology and Aging, **17(1)**, pp. 101.

[16] Allport GW, 1954, *The nature of prejudice*, Addison-Wesley, Boston (MA).

[17] Amer M, Daim TU & Jetter A, 2013, *A review of scenario planning*, Futures, **46**, pp. 23–40.

[18] AMEU, 2012, *Association of municipal electricity utilities*, [Online], [Cited October 2019], Available from http://www.ameu.co.za/.

[19] Andersen SM & Glassman NS, 1996, *Responding to significant others when they are not there: Effects on interpersonal inference, motivation, and affect*, pp. 262–321 in Sorrentino R & Higgins E (Eds), *Handbook of motivation and cognition*, Guilford Press, New York (NY).

[20] Anderson BH & Maletta M, 1994, *Auditor attendance to negative and positive information: The effect of experience-related differences*, Behavioral Research in Accounting, **6**, pp. 1–20.

[21] Anderson JC, Jennings MM, Lowe DJ & Reckers PM, 1997, *The mitigation of hindsight bias in judges' evaluation of auditor decisions*, Auditing, **16(2)**, pp. 20–39.

[22] Areff A, 2015, *Zero increase in fees — Zuma*, [Online], [Cited November 2018], Available from https://www.news24.com/SouthAfrica/News/Zero-increase-in-fees-Zuma-20151023.

[23] AREI, 2015, *About AREI*, [Online], [Cited October 2019], Available from http://www.arei.org/#about.

[24] Arkes HR, Wortmann RL, Saville PD & Harkness AR, 1981, *Hindsight bias among physicians weighing the likelihood of diagnoses*, Journal of Applied Psychology, **66(2)**, pp. 252–254.

[25] Arkes HR, 1991, *Costs and benefits of judgment errors: Implications for debiasing*, Psychological Bulletin, **110(3)**, pp. 486–452.

[26] Arkes HR & Ayton P, 1999, *The sunk cost and Concorde effects: Are humans less rational than lower animals*, Psychological Bulletin, **125(5)**, pp. 591–600.

[27] Arkes HR, Faust D, Guilmette TJ & Hart K, 1988, *Eliminating the hindsight bias*, Journal of Applied Psychology, **73(2)**, pp. 305–307.

[28] Arlin PK, 1989, *The problem of the problem*, pp. 229–237 in Sinnot JD (Ed), *Everyday problem solving: Theory and applications*, Praeger, New York (NY).

[29] Aron A, Aron EN & Smollan D, 1992, *Inclusion of other in the self scale and the structure of interpersonal closeness*, Journal of Personality and Social Psychology, **63(4)**, pp. 596–612.

[30] Aron A, Aron EN, Tudor M & Nelson G, 2004, *Close relationships as including other in the self*, pp. 365–379 in Reis HT & Rusbult CE (Eds), *Close relationships: Key readings*, Psychology Press, London.

[31] Aronson E, 1969, *The theory of cognitive dissonance: A current perspective*, Advances in Experimental Social Psychology, **4**, pp. 1–34.

[32] Asch SE, 1951, *Effects of group pressure upon the modification and distortion of judgments*, pp. 177–190 in Guetzkow H (Ed), *Groups, leadership and men; research in human relations*, Carnegie Press, Oxford.

[33] Ashby WR, 1961, *An introduction to cybernetics*, Chapman & Hall Ltd, London.

[34] Ashby WR & Goldstein J, 2011, *Variety, constraint, and the law of requisite variety*, Emergence: Complexity and Organization, **13(1/2)**, pp. 190–197.

[35] Ashraf N, Karlan D & Yin W, 2006, *Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines*, Quarterly Journal of Economics, **121(2)**, pp. 635–672.

[36] AZ Quotes, 2015, *Robert Gates quote*, [Online], [Cited May 2018], Available from http://www.azquotes.com/quote/992597.

[37] Azar A, Khosravani F & Jalali R, 2014, *Drama theory: A problem structuring method in soft OR (A practical application: Nuclear negotiations analysis between Islamic Republic of Iran and the 5+1 group)*, International Journal of Humanities, **19(4)**, pp. 1–14.

[38] Babcock L & Loewenstein G, 1997, *Explaining bargaining impasse: The role of self-serving biases*, Journal of Economic Perspectives, **11(1)**, pp. 109–126.

[39] Baker S, Young S, Kousis M & Richardson D, 1997, *The politics of sustainable development: Theory, policy and practice within the European Union*, Psychology Press, London.

[40] Bales RF, 1951, *Interaction process analysis: A method for the study of small groups*, Social Forces, **29(3)**, pp. 329–330.

[41] Bales RF, 2000, *Social interaction systems: Theory and measurement — Book review*, Group Dynamics Theory Research and Practice, **4(2)**, pp. 199–208.

[42] Banks J & Rossini FA, 1987, *Management science failures in the public sector*, Public Productivity Review, **10(4)**, pp. 15–23.

[43] Bar-Hillel M, 1979, *The role of sample size in sample evaluation*, Organizational Behavior and Human Performance, **24(2)**, pp. 245–257.

[44] Bargh JA, 1989, *Conditional automaticity: Varieties of automatic influence in social perception and cognition*, Unintended Thought, **3**, pp. 51–69.

[45] Bargh JA, 1997, *Automaticity of everyday life*, pp. 1–61 in Wyer R (Ed), *Advances in social cognition*, Lawrence Erlbaum Associates, Mahwah (NJ).

[46] Barker E, 2012, *The political thought of Plato and Aristotle*, Courier Corporation, Chelmsford (MA).

[47] Baron J & Hershey JC, 1988, *Outcome bias in decision evaluation*, Journal of Personality and Social Psychology, **54(4)**, pp. 569–579.

[48] Baron J & Ritov I, 2004, *Omission bias, individual differences, and normality*, Organizational Behavior and Human Decision Processes, **94(2)**, pp. 74–85.

[49] Bassok M, 2003, *Analogical transfer in problem solving*, pp. 343–369 in Davidson J (Ed), *The psychology of problem solving*, Cambridge University Press, New York (NY).

[50] Baumeister RF & Leary MR, 1995, *The need to belong: Desire for interpersonal attachments as a fundamental human motivation*, Psychological Bulletin, **117(3)**, pp. 497–529.

[51] Bazerman MH & Moore DA, 2008, *Judgment in managerial decision making*, 8th Edition, Wiley, New York (NY).

[52] Beer S, 1960, *Cybernetics and management*, English Universities Press, London.

[53] Beer S, 1966, *Decision and control: The meaning of operational research and management cybernetics*, John Wiley & Sons, New York (NY).

[54] Beer S, 1979, *The heart of enterprise*, 2nd Edition, Wiley, Chichester.

[55] Beer S, 2002, *What is cybernetics?*, Kybernetes, **31(2)**, pp. 209–219.

[56] Beishon RJ & Peters G, 1976, *Systems behaviour*, Open University Press, Maidenhead.

[57] Belton V & Stewart T, 2010, *Problem structuring and multiple criteria decision analysis*, pp. 209–239 in Ehrgott M, Figueira JR & Greco S (Eds), *Trends in multiple criteria decision analysis*, Springer, Berlin.

[58] Bem DJ, 1972, *Self-perception theory*, Advances in Experimental Social Psychology, **6**, pp. 1–62.

[59] Benjamin D, Moore D & Rabin M, 2013, *Misconceptions of chance: Evidence from an integrated experiment*, [Online], [Cited September 2017], Available from https://scholar.harvard.edu/files/rabin/files/misconceptions.pdf.

[60] Berkeley Lab, 2019, *About microgrids*, [Online], [Cited October 2019], Available from https://building-microgrid.lbl.gov/about-microgrids.

[61] Berkman LF, Leo-Summers L & Horwitz RI, 1992, *Emotional support and survival after myocardial infarction*, Annals of Internal Medicine, **117(12)**, pp. 1003–1009.

[62] Bernstein PL, 1996, *Against the gods: The remarkable story of risk*, Wiley, New York (NY).

[63] Beshears J, Choi J, Laibson D & Madrian B, 2019, *Active choice and health care costs: Evidence from prescription drug home delivery*, [Online], [Cited October 2012], Available from https://scholar.harvard.edu/bmadrian/working-papers.

[64] Biernacki P & Waldorf D, 1981, *Snowball sampling: Problems and techniques of chain referral sampling*, Sociological Methods & Research, **10(2)**, pp. 141–163.

[65] Bishop CM, 2006, *Pattern recognition and machine learning*, Springer, New York (NY).

[66] Blanton H, Pelham BW, DeHart T & Carvallo M, 2001, *Overconfidence as dissonance reduction*, Journal of Experimental Social Psychology, **37(5)**, pp. 373–385.

[67] Blascovich J, Ginsburg GP & Howe RC, 1976, *Blackjack, choice shifts in the field*, Sociometry, **39(3)**, pp. 274–276.

[68] Bornstein BH & Emler AC, 2001, *Rationality in medical decision making: A review of the literature on doctor's decision-making biases*, Journal of Evaluation in Clinical Practice, **7(2)**, pp. 97–107.

[69] Bourne L, 2005, *Project relationship management and the stakeholder circle*, PhD thesis, Royal Melbourne Institute of Technology, Melbourne.

[70] Brainy Quote, 2001, *Karl Marx quotes*, [Online], [Cited October 2018], Available from https://www.brainyquote.com/quotes/karl_marx_382655.

[71] Brewer MB & Campbell DT, 1976, *Ethnocentrism and intergroup attitudes: East African evidence*, Sage, New York (NY).

[72] Brewer MB & Gardner W, 1996, *Who is this We? Levels of collective identity and self representations*, Journal of Personality and Social Psychology, **71(1)**, pp. 83–93.

[73] Brink vdS, 2019, *Corvus*, [Online], [Cited October 2015], Available from http://www.corvus.co.za/profile.htm.

[74] Brown CL, 1999, *"Do the right thing:" Diverging effects of accountability in a managerial context*, Marketing Science, **18(3)**, pp. 230–246.

[75] Browne K, 2005, *Snowball sampling: Using social networks to research non-heterosexual women*, International Journal of Social Research Methodology, **8(1)**, pp. 47–60.

[76] Brugha R & Varvasovszky Z, 2000, *Stakeholder analysis: A review*, Health Policy and Planning, **15(3)**, pp. 239–246.

[77] Buehler R, Griffin D & Ross M, 1994, *Exploring the planning fallacy: Why people underestimate their task completion times*, Journal of Personality and Social Psychology, **67(3)**, pp. 366–381.

[78] Bugaje IM, 2006, *Renewable energy for sustainable development in Africa: A review*, Renewable and Sustainable Energy Reviews, **10(6)**, pp. 603–612.

[79] Burge J & McCall R, 2015, *Diagnosing wicked problems*, pp. 313–326 in Gero J & Hanna S (Eds), *Design computing and cognition*, Springer, New York (NY).

[80] Burnstein E & Vinokur A, 1973, *Testing two classes of theories about group induced shifts in individual choice*, Journal of Experimental Social Psychology, **9(2)**, pp. 123–137.

[81] Burnstein E & Vinokur A, 1977, *Persuasive argumentation and social comparison as determinants of attitude polarization*, Journal of Experimental Social Psychology, **13(4)**, pp. 315–332.

[82] Burrell G & Morgan G, 2017, *Sociological paradigms and organisational analysis: Elements of the sociology of corporate life*, Routledge, London.

[83] Cacioppo JT, 2002, *Foundations in social neuroscience*, MIT Press, Cambridge (MA).

[84] Caluza N, 2017, *Does corruption reign in the department of education?*, [Online], [Cited November 2018], Available from http://www.sadtu.org.za/content/does-corruption-reign-department-education.

[85] Camerer CF & Hogarth RM, 1999, *The effects of financial incentives in experiments: A review and capital-labor-production framework*, Journal of Risk and Uncertainty, **19(1-3)**, pp. 7–42.

[86] Campbell DT, 1958, *Common fate, similarity, and other indices of the status of aggregates of persons as social entities*, Systems Research and Behavioral Science, **3(1)**, pp. 14–25.

[87] Capra F, 2007, *Complexity and life*, Systems Research and Behavioral Science, **24(5)**, pp. 475–479.

[88] Carbonell JG, Michalski RS & Mitchell TM, 1983, *An overview of machine learning*, pp. 3–23 in Michalski RS, Carbonell JG & Mitchell TM (Eds), *Machine learning, Volume I*, Elsevier, Amsterdam.

[89] Carretie L, Mercado F, Tapia M & Hinojosa JA, 2001, *Emotion, attention, and the negativity bias, studied through event-related potentials*, International Journal of Psychophysiology, **41(1)**, pp. 75–85.

[90] Carrol GD, Choi JJ, Laibson D, C MB & A M, 2009, *Optimal defaults and active decisions*, Quarterly Journal of Economics, **124**, pp. 1639–1674.

[91] Carroll AB, 1991, *The pyramid of corporate social responsibility: Toward the moral management of organizational stakeholders*, Business Horizons, **34(4)**, pp. 39–48.

[92] Carter-Brown C & Wright J, 2018, *Municipal energy transition — Opportunities for new business models and revenue streams*, [Online], [Cited October 2019], Available from https://www.salga.org.za/SALGA%20Municipal%20Managers%20Forum%20Web/Documents/Day%201%20MMF%20Mpumalanga/CSIR%2020%20Feb%202018.pdf.

[93] Carter CR, Kaufmann L & Michel A, 2007, *Behavioral supply management: A taxonomy of judgment and decision-making biases*, International Journal of Physical Distribution and Logistics Management, **37(8)**, pp. 631–669.

[94] Cartwright DE & Zander AE, 1953, *Group dynamics research and theory*, Harper and Row, New York (NY).

[95] Cervone D & Peake PK, 1986, *Anchoring, efficacy, and action: The influence of judgmental heuristics on self-efficacy judgments and behavior*, Journal of Personality and Social Psychology, **50(3)**, pp. 492–501.

[96] Chapelle O, Scholkopf B & Zien A, 2009, *Semi-supervised learning*, IEEE Transactions on Neural Networks, **20(3)**, pp. 542–542.

[97] Chapman LJ & Chapman JP, 1967, *Genesis of popular but erroneous psychodiagnostic observations*, Journal of Abnormal Psychology, **72(3)**, pp. 193–204.

[98] Checkland PB, 1975, *The development of systems thinking*, Progress in Cybernetics and Systems Research, **2**, pp. 278–283.

[99] Checkland PB, 1978, *The origins and nature of hard systems thinking*, Journal of Applied Analysis, **5(2)**, pp. 99.

[100] Checkland PB, 1981, *Systems thinking, systems practice*, John Wiley, Chichester.

[101] Checkland PB, 1999, *Systems thinking*, pp. 45–56 in Currie W & Galliers B (Eds), *Rethinking management information systems*, Oxford University Press, Oxford.

[102] Checkland PB, 2000, *Soft systems methodology: A thirty year retrospective*, Systems Research and Behavioral Science, **17(1)**, pp. 11–58.

[103] Checkland PB & Poulter J, 2010, *Soft systems methodology*, pp. 191–242 in Reynolds M & Holwell S (Eds), *Systems approaches to managing change: A practical guide*, Springer, London.

[104] Cheema A & Soman D, 2008, *The effect of partitions on controlling consumption*, Journal of Marketing Research, **45(6)**, pp. 665–675.

[105] Chen S, Boucher HC & Tapias MP, 2006, *The relational self revealed: Integrative conceptualization and implications for interpersonal life*, Psychological Bulletin, **132(2)**, pp. 151–179.

[106] Chen S, Boucher H & Kraus MW, 2011, *The relational self*, pp. 149–175 in Schwartz SJ, Luyckx K & Vignoles VL (Eds), *Handbook of identity theory and research*, Springer, New York (NY).

[107] Chi MTH, Feltovich PJ & Glaser R, 1981, *Categorization and representation of physics problems by experts and novices*, Cognitive Science, **5(2)**, pp. 121–152.

[108] Child J, 1997, *Strategic choice in the analysis of action, structure, organizations and environment: Retrospect and prospect*, Organization Studies, **18(1)**, pp. 43–76.

[109] Chohan UW, 2017, *Initial coin offerings: Risks, regulation, and accountability*, [Online], [Cited October 2019], Available from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3080098.

[110] Christensen-Szalanski JJJ & Willham CF, 1991, *The hindsight bias: A meta-analysis*, Organizational Behavior and Human Decision Processes, **48(1)**, pp. 147–168.

[111] Chugh D, 2004, *Societal and managerial implications of implicit social cognition: Why milliseconds matter*, Social Justice Research, **17(2)**, pp. 203–222.

[112] Churchman CW, 1970, *Operations research as a profession*, Management Science, **17(2)**, pp. 37–53.

[113] Churchman CW, 1971, *The design of inquiring systems: Basic concepts of systems and organization*, Basic Books, New York (NY).

[114] Churchman CW, 1979, *The systems approach and its enemies*, Basic Books, New York (NY).

[115] Churchman CW, Ackoff RL & Arnoff EL, 1957, *Introduction to operations research*, Wiley, New York (NY).

[116] City of Cape Town, 2018, *City of Cape Town clean energy vision*, [Online], [Cited October 2019], Available from https://sanea.org.za/wp-content/uploads/2018/07/SANEA-Energy-Rendezvous-CPT-25-July-2018-Kadri-Nassiep-Clean-Energy-Vision-for-CoCT.pdf.

[117] Clarke M & Keys P, 1981, *What happens when public sector employees are made redundant: Some exploratory results using a micro-simulation model*, Unpublished Report, School of Geography, University of Leeds, Leeds.

[118] Confrey J, 1991, *Learning to listen: A student's understanding of powers of ten*, Radical Constructivism in Mathematics Education, **1(1)**, pp. 111–138.

[119] Constitutional Assembly, Republic of South Africa, 1996, *Constitution of the Republic of South Africa 1996: As adopted by the Constitutional Assembly on 8 May 1996*, Government Printer, Pretoria.

[120] Craik FI & McDowd JM, 1987, *Age differences in recall and recognition*, Journal of Experimental Psychology: Learning, Memory, and Cognition, **13(3)**, pp. 474.

[121] Crocker J & Luhtanen R, 1990, *Collective self-esteem and ingroup bias*, Journal of Personality and Social Psychology, **58(1)**, pp. 60–67.

[122] Cunningham C & Tynan C, 1990, *Electronic trading and retail supply chain relationships: A network approach*, Unpublished Report, Henley Management College, Reading.

[123] Cyert RM & March JG, 1963, *A behavioral theory of the firm*, Prentice Hall, Englewood Cliffs (NJ).

[124] de Moll KE, 2010, *Everyday experiences of power*, PhD thesis, University of Tennessee, Knoxville (TN).

[125] Deetz S, 1996, *Crossroads — Describing differences in approaches to organization science: Rethinking Burrell and Morgan and their legacy*, Organization Science, **7(2)**, pp. 191–207.

[126] Defeyter MA, Russo R & McPartlin PL, 2009, *The picture superiority effect in recognition memory: A developmental study using the response signal procedure*, Cognitive Development, **24(3)**, pp. 265–273.

[127] Dennis B, 2013, *Qualitative research: A reader in philosophy, core concepts, and practice*, Peter Lang, Bern.

[128] Department of Energy, 2010, *Corporate profile*, [Online], [Cited October 2019], Available from http://www.energy.gov.za/files/au_frame.html.

[129] Department of Higher Education and Training, 2013, *About us*, [Online], [Cited October 2018], Available from http://www.dhet.gov.za/SitePages/AboutUsNew.aspx.

[130] Deutsch M & Gerard HB, 1955, *A study of normative and informational social influences upon individual judgment*, Journal of Abnormal and Social Psychology, **51(3)**, pp. 629.

[131] Dholakia UM & Bagozzi R, 2003, *As time goes by: How goal and implementation intentions influence enactment of short-fuse behaviors*, Journal of Applied Social Psychology, **33(5)**, pp. 889–922.

[132] Dion KL, 2000, *Group cohesion: From field of forces to multidimensional construct*, Group Dynamics: Theory, Research, and Practice, **4(1)**, pp. 7–26.

[133] Doise W, 1969, *Intergroup relations and polarization of individual and collective judgments*, Journal of Personality and Social Psychology, **12(2)**, pp. 136–143.

[134] Doyle J, 1999, *Rational decision making*, pp. 701–703, *MIT Encyclopedia of the Cognitive Sciences*, MIT Press, Cambridge (MA).

[135] Dreu CKW, Nauta A & Vliert E, 1995, *Self-serving evaluations of conflict behavior and escalation of the dispute*, Journal of Applied Social Psychology, **25(23)**, pp. 2049–2066.

[136] Drever E, 1995, *Using semi-structured interviews in small-scale research*, Scottish Council for Research in Education, Edinburgh.

[137] Dror IE & Cole SA, 2010, *The vision in 'blind' justice: Expert perception, judgment, and visual cognition in forensic pattern recognition*, Psychonomic Bulletin and Review, **17(2)**, pp. 161–167.

[138] Duncan J, 2016, *Why student protests turn violent*, [Online], [Cited October 2018], Available from https://www.iol.co.za/dailynews/opinion/why-student-protests-turn-violent-2074612.

[139] Dunning D, Griffin DW, Milojkovic JD & Ross L, 1990, *The overconfidence effect in social prediction*, Journal of Personality and Social Psychology, **58(4)**, pp. 568–581.

[140] Durand R, 2003, *Predicting a firm's forecasting ability: The roles of organizational illusion of control and organizational attention*, Strategic Management Journal, **24(9)**, pp. 821–838.

[141] Eden C & Ackermann F, 1998, *Making strategy: The journey of strategic management*, Sage Publications, London.

[142] Epley N & Gilovich T, 2001, *Putting adjustment back in the anchoring and adjustment heuristic: Differential processing of self-generated and experimenter-provided anchors*, Psychological Science, **12(5)**, pp. 391–396.

[143] Epley N & Gilovich T, 2006, *The anchoring-and-adjustment heuristic: Why the adjustments are insufficient*, Psychological Science, **17(4)**, pp. 311–318.

[144] Epley N, Keysar B, Van Boven L & Gilovich T, 2004, *Perspective taking as egocentric anchoring and adjustment*, Journal of Personality and Social Psychology, **87(3)**, pp. 327–339.

[145] Epstein S, 1994, *Integration of the cognitive and the psychodynamic unconscious*, American Psychologist, **49(8)**, pp. 709–724.

[146] Esser JK & Lindoerfer JS, 1989, *Groupthink and the space shuttle Challenger accident: Toward a quantitative case analysis*, Journal of Behavioral Decision Making, **2(3)**, pp. 167–177.

[147] Etikan I, Musa SA & Alkassim RS, 2016, *Comparison of convenience sampling and purposive sampling*, American Journal of Theoretical and Applied Statistics, **5(1)**, pp. 1–4.

[148] Etzioni A, 1964, *Modern organizations*, Prentce Hall, Englewood Cliffs (NJ).

[149] Evans JBT, 2003, *In two minds: Dual-process accounts of reasoning*, Trends in Cognitive Sciences, **7(10)**, pp. 454–459.

[150] Evans JBT & Curtis-Holmes J, 2005, *Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning*, Thinking and Reasoning, **11(4)**, pp. 382–389.

[151] Evans JBT & Dusoir AE, 1977, *Proportionality and sample size as factors in intuitive statistical judgement*, Acta Psychologica, **41(2)**, pp. 129–137.

[152] Eyewitness News, 2016, *Opinion: Why student protests in SA have turned violent*, [Online], [Cited October 2018], Available from https://ewn.co.za/2016/09/30/OPINION-Why-student-protests-in-South-Africa-have-turned-violent.

[153] Fagerlin A, Wang C & Ubel PA, 2005, *Reducing the influence of anecdotal reasoning on people's health care decisions: Is a picture worth a thousand statistics?*, Medical Decision Making, **25(4)**, pp. 398–405.

[154] Fagley NS & Miller PM, 1990, *The effect of framing on choice: Interactions with risk-taking propensity, cognitive style, and sex*, Personality and Social Psychology Bulletin, **16(3)**, pp. 496–510.

[155] Farrokhi F & Mahmoudi-Hamidabad A, 2012, *Rethinking convenience sampling: Defining quality criteria*, Theory and Practice in Language Studies, **2(4)**, pp. 647–656.

[156] Ferraris C & Carveth R, 2003, *NASA and the Columbia disaster: Decision-making by groupthink?*, Proceedings of the 2003 Association for Business Communication Annual Convention, Albuquerque (NM), pp. 1–12.

[157] Festinger L, 1962, *A theory of cognitive dissonance*, Stanford University Press, Stanford (CA).

[158] Fetherstonhaugh D, Slovic P, Johnson S & Friedrich J, 1997, *Insensitivity to the value of human life: A study of psychophysical numbing*, Journal of Risk and Uncertainty, **14(3)**, pp. 283–300.

[159] Fischhoff B, 1981, *Debiasing*, pp. 422–444 in Kahneman D, Slovic P & Tversky A (Eds), *Judgment under uncertainty: Heuristics and biases*, Cambridge University Press, New York (NY).

[160] Fiske ST, 2004, *What's in a category? Responsibility, intent, and the avoidability of bias against outgroups*, pp. 127–140 in Miller A (Ed), *The social psychology of good and evil*, Guilford Press, New York (NY).

[161] Flood RL, 1991, *Implementing total quality management through total systems intervention: A creative approach to problem solving*, pp. 43–59 in Blackham RL, Flood RL, Jackson M, Mansell GJ & Probert SVE (Eds), *Systems thinking in Europe*, Springer, New York (NY).

[162] Flood RL, 1995, *Solving problem solving*, Wiley, Chichester.

[163] Flood RL, 1995, *Total systems intervention (TSI): A reconstitution*, Journal of the Operational Research Society, **46(2)**, pp. 174–191.

[164] Flood RL & Jackson MC, 1991, *Critical systems thinking*, Springer, New York (NY).

[165] Flood RL & Jackson MC, 1991, *Total systems intervention: A practical face to critical systems thinking*, Systems Practice, **4(3)**, pp. 197–213.

[166] Fong GT, Krantz DH & Nisbett RE, 1986, *The effects of statistical training on thinking about everyday problems*, Cognitive Psychology, **18(3)**, pp. 253–292.

[167] Fong GT & Nisbett RE, 1991, *Immediate and delayed transfer of training effects in statistical reasoning*, Journal of Experimental Psychology, **120(1)**, pp. 34–35.

[168] Forrester JW, 1994, *System dynamics, systems thinking, and soft OR*, System Dynamics Review, **10(2-3)**, pp. 245–256.

[169] Forsyth DR, 2008, *The power of groups*, The Group Psychologist, **18(1)**, pp. 4–22.

[170] Forsyth DR, 2009, *Group dynamics*, 5th Edition, Cengage Learning, Boston (MA).

[171] Frazer JG, 1959, *The golden bough: A study in magic and religion*, Macmillan Press Ltd, London.

[172] Freeman RE, 1984, *Strategic management: A stakeholder approach*, Pitman, Boston (MA).

[173] Freeman T, 2006, *Best practice in focus group research: Making sense of different views*, Journal of Advanced Nursing, **56(5)**, pp. 491–497.

[174] French J & Raven B, 1959, *The bases of social power*, pp. 251–260 in Sharitz JM, Ott JS & Jang YS (Eds), *Classics of organisation theory*, Cengage Learning, Boston (MA).

[175] Freud S, 1958, *The dynamics of transference*, Hogarth Press, London.

[176] Friedkin NE, 1999, *Choice shift and group polarization*, American Sociological Review, **37**, pp. 856–875.

[177] Friend J, 2011, *The strategic choice approach*, pp. 22–45 in Keskinocak P, Kharoufeh JP & Smith JC (Eds), *Wiley encyclopedia of operations research and management science*, Wiley, Hoboken (NJ).

[178] Friend J & Hickling A, 2005, *Planning under pressure: The strategic choice approach*, Routledge, Oxford.

[179] Galbraith JK, 1983, *The anatomy of power*, Challenge, **26(3)**, pp. 26–33.

[180] Galesic M, Garcia-Retamero R & Gigerenzer G, 2009, *Using icon arrays to communicate medical risks: Overcoming low numeracy*, Health Psychology, **28(2)**, pp. 210–216.

[181] Galton F, 1886, *Regression towards mediocrity in hereditary stature*, Journal of the Anthropological Institute of Great Britain and Ireland, **15**, pp. 246–263.

[182] Galton F, 1907, *Vox populi (the wisdom of crowds)*, Nature, **75(7)**, pp. 450–451.

[183] Garb HN, 1996, *The representativeness and past-behavior heuristics in clinical judgment*, Professional Psychology: Research and Practice, **27(3)**, pp. 272–277.

[184] Gauteng Province Department of Education, 2017, *About education*, [Online], [Cited October 2018], Available from `http://www.gauteng.gov.za/government/departments/education/Pages/index.aspx`.

[185] Gentleman R & Carey VJ, 2008, *Unsupervised machine learning*, pp. 137–157 in Hahne F, Gentleman R, Carey VJ & Falcon S (Eds), *Bioconductor case studies*, Springer, New York (NY).

[186] Gigerenzer G & Todd PM, 1999, *Simple heuristics that make us smart*, Oxford University Press, Oxford.

[187] Gigerenzer G & Hoffrage U, 1995, *How to improve Bayesian reasoning without instruction: Frequency formats*, Psychological Review, **102(4)**, pp. 684–704.

[188] Gilovich T, Griffin D & Kahneman D, 2002, *Heuristics and biases: The psychology of intuitive judgment*, Cambridge University Press, Cambridge.

[189] Goethals GR & Zanna MP, 1979, *The role of social comparison in choice shifts*, Journal of Personality and Social Psychology, **37(9)**, pp. 1469–1476.

[190] Goldstein DG, 1997, *Models of bounded rationality for inference*, PhD thesis, University of Chicago, Chicago (IL).

[191] Goldstein DG & Gigerenzer G, 2002, *Models of ecological rationality: The recognition heuristic*, Psychological Review, **109(1)**, pp. 75–90.

[192] Gollwitzer PM & Sheeran P, 2006, *Implementation intentions and goal achievement: A meta-analysis of effects and processes*, Advances in Experimental Social Psychology, **38**, pp. 69–119.

[193] Gouran DS, Hirokawa RY & Martz AE, 1986, *A critical analysis of factors related to decisional processes involved in the Challenger disaster*, Communication Studies, **37(3)**, pp. 118–135.

[194] GreenCape, 2019, *Market intelligence report*, [Online], [Cited October 2019], Available from https://www.greencape.co.za/assets/Uploads/ENERGY-SERVICES-MARKET-INTELLIGENCE-REPORT-FINAL-WEB.pdf.

[195] Grobbelaar SS, 2019, *ResearchGate*, [Online], [Cited October 2019], Available from https://www.researchgate.net/profile/Sara_Grobbelaar.

[196] Guba EG & Lincoln YS, 1994, *Competing paradigms in qualitative research*, pp. 105–117 in Denzin NK & Lincoln YS (Eds), *Handbook of qualitative research*, Sage Publications, Oaks (CA).

[197] Haizlip NJ M, Schorling J, Williams A & Plews-Ogan M, 2012, *Perspective: The negativity bias, medical education, and the culture of academic medicine: Why culture change is hard*, Academic Medicine, **87(9)**, pp. 1205–1209.

[198] Hall AD, 1962, *A methodology for systems engineering*, Van Nostrand, Princeton (NJ).

[199] Hamann R, 2003, *Mining companies' role in sustainable development: The 'why' and 'how' of corporate social responsibility from a business perspective*, Development Southern Africa, **20(2)**, pp. 237–254.

[200] Hammond JS, Keeney RL & Raiffa H, 1999, *Smart choices: A practical guide to making better life decisions*, Crown Publishing Group, New York (NY).

[201] Hammond KR, 2000, *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*, Oxford University Press, Oxford.

[202] Hardman D & Hardman DK, 2009, *Judgment and decision making: Psychological perspectives*, John Wiley & Sons, Hoboken (NJ).

[203] Harris DM & Guten S, 1979, *Health-protective behavior: An exploratory study*, Journal of Health and Social Behavior, **20(1)**, pp. 17–29.

[204] Harvey JB, 1996, *The Abilene paradox and other meditations on management*, Lexington Books, Lexington (MA).

[205] Haselton MG, Nettle D & Murray DR, 2005, *The evolution of cognitive bias*, pp. 968–987 in Buss MB (Ed), *The handbook of evolutionary psychology*, John Wiley & Sons Incorporated, Hoboken (NJ).

[206] Haslam SA, 2004, *Psychology in organizations*, 2nd Edition, Sage, London.

[207] Hayes JR, 2013, *The complete problem solver*, 2nd Edition, Routledge, London.

[208] Head BW, 2008, *Wicked problems in public policy*, Public Policy, **3(2)**, pp. 101–118.

[209] Healy MJR & Goldstein H, 1978, *Regression to the mean*, Annals of Human Biology, **5(3)**, pp. 277–280.

[210] Herbert TT & Estes RW, 1977, *Improving executive decisions by formalizing dissent: The corporate devil's advocate*, Academy of Management Review, **2(4)**, pp. 662–667.

[211] Herman H, 2019, *A framework for the design, development and implementation of technology platforms in the South African health context*, MEng thesis, Stellenbosch University, Stellenbosch.

[212] Herzberg F, 1968, *One more time: How do you motivate employees*, Harvard Business Review, **46(1)**, pp. 53–62.

[213] Hewstone M, 1990, *The ultimate attribution error? A review of the literature on intergroup causal attribution*, European Journal of Social Psychology, **20(4)**, pp. 311–335.

[214] Hewstone M, Rubin M & Willis H, 2002, *Intergroup bias*, Annual Review of Psychology, **53(1)**, pp. 575–604.

[215] Higgins ET, 1989, *Knowledge accessibility and activation: Subjectivity and suffering from unconscious sources*, Unintended Thought, **3**, pp. 75–123.

[216] Hoffrage U, Lindsey S, Hertwig R & Gigerenzer G, 2000, *Communicating statistical information*, Science, **290**, pp. 2261–2262.

[217] Hogarth RM & Einhorn HJ, 1992, *Order effects in belief updating: The belief-adjustment model*, Cognitive Psychology, **24(1)**, pp. 1–55.

[218] Hogg MA, 2001, *Social categorization, depersonalization, and group behavior*, pp. 56–85 in Hogg MA & Tindale RS (Eds), *Blackwell handbook of social psychology: Group processes*, Blackwell Publishers Limited, Hoboken (NJ).

[219] Hogg MA, 2016, *Social identity theory*, pp. 3–17 in McKeown S, Haji R & Ferguson N (Eds), *Understanding peace and conflict through social identity theory*, Springer, New York (NY).

[220] Hogg MA & Terry DI, 2000, *Social identity and self-categorization processes in organizational contexts*, Academy of Management Review, **25(1)**, pp. 121–140.

[221] Honderich T, 2005, *The Oxford companion to philosophy*, Oxford University Press, Oxford.

[222] Hopwood B, Mellor M & O'Brien G, 2005, *Sustainable development: Mapping different approaches*, Sustainable Development, **13(1)**, pp. 38–52.

[223] Hornsey MJ & Jetten J, 2004, *The individual within the group: Balancing the need to belong with the need to be different*, Personality and Social Psychology Review, **8(3)**, pp. 248–264.

[224] Howard N, 1999, *Confrontation analysis: How to win operations other than war*, Command and Control Research Portal, Washington (DC).

[225] Hsee CK & Zhang J, 2004, *Distinction bias: Misprediction and mischoice due to joint evaluation*, Journal of Personality and Social Psychology, **86(5)**, pp. 680–695.

[226] Hubert H & Mauss M, 1902, *A general theory of magic*, Norton, New York (NY).

[227] Insko CA, Pinkley RL, Hoyle RH, Dalton B, Hong G, Slim RM, Landry P, Holton B, Ruffin PF & Thibaut J, 1987, *Individual versus group discontinuity: The role of intergroup contact*, Journal of Experimental Social Psychology, **23(3)**, pp. 250–267.

[228] Jabareen Y, 2009, *Building a conceptual framework: Philosophy, definitions, and procedure*, International Journal of Qualitative Methods, **8(4)**, pp. 49–62.

[229] Jackson MC & Keys P, 1984, *Towards a system of systems methodologies*, Journal of the Operational Research Society, **35(6)**, pp. 473–486.

[230] Jackson MC, 1985, *Social systems theory and practice: The need for a critical approach*, International Journal of General Systems, **10(2-3)**, pp. 135–151.

[231] Jackson MC, 1990, *Beyond a system of systems methodologies*, Journal of the Operational Research Society, **41(8)**, pp. 657–668.

[232] Jackson MC, 1992, *Systems methodology for the management sciences*, Springer Science & Business Media, New York (NY).

[233] Jackson MC, 2003, *Systems thinking: Creative holism for managers*, Wiley, Chichester.

[234] Jackson MC, 2007, *Systems approaches to management*, Springer Science & Business Media, New York (NY).

[235] Jacoby LL & Dallas M, 1981, *On the relationship between autobiographical memory and perceptual learning*, Journal of Experimental Psychology: General, **110(3)**, pp. 306–340.

[236] Jacowitz KE & Kahneman D, 1995, *Measures of anchoring in estimation tasks*, Personality and Social Psychology Bulletin, **21(11)**, pp. 1161–1166.

[237] Jain S & Jain P, 2017, *The rise of renewable energy implementation in South Africa*, Energy Procedia, **143**, pp. 721–726.

[238] James W, 1890, *The consciousness of self*, Principles of Psychology, **8**, pp. 291–402.

[239] Janis IL, 1971, *Groupthink*, Psychology Today, **5(6)**, pp. 43–46.

[240] Janis IL, 1972, *Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes*, Houghton Mifflin, Boston (MA).

[241] Janis IL, 1982, *Groupthink: Psychological studies of policy decisions and fiascoes*, Houghton Mifflin, Boston (MA).

[242] Jepsen AL & Eskerod P, 2009, *Stakeholder analysis in projects: Challenges in using current guidelines in the real world*, International Journal of Project Management, **27(4)**, pp. 335–343.

[243] Jepson C, Krantz DH & Nisbett RE, 1983, *Inductive reasoning: Competence or skill*, Behavioral and Brain Sciences, **6(3)**, pp. 494–501.

[244] Johnson EJ, Bellman S & Lohse GL, 2002, *Defaults, framing and privacy: Why opting in-opting out*, Marketing Letters, **13(1)**, pp. 5–15.

[245] Johnson EJ & Goldstein D, 2003, *Do defaults save lives?*, Science, pp. 1338–1339.

[246] Johnson MP & Smilowitz K, 2012, *Community-based operations research*, pp. 37–65 in Johnson M (Ed), *Community-based operations research*, Springer, New York (NY).

[247] Jonassen DH, 1997, *Instructional design models for well-structured and Ill-structured problem-solving learning outcomes*, Educational Technology Research and Development, **45(1)**, pp. 65–94.

[248] Jonassen DH, 2000, *Toward a design theory of problem solving*, Educational Technology Research and Development, **48(4)**, pp. 63–85.

[249] Jonassen DH, 2001, *Can you train employees to solve problems?*, Performance Improvement, **40(9)**, pp. 18–24.

[250] Jonassen DH, 2010, *Learning to solve problems: A handbook for designing problem-solving learning environments*, Routledge, London.

[251] Jonassen DH & Hung W, 2008, *All problems are not equal: Implications for problem-based learning*, Interdisciplinary Journal of Problem-Based Learning, **2(2)**, pp. 4–28.

[252] Joubert AR, Janssen R & Stewart TJ, 2008, *Allocating fishing rights in South Africa: A participatory approach*, Fisheries Management and Ecology, **15(1)**, pp. 27–37.

[253] Joubert A, Stewart TJ, Janssen R & Gilbert AJ, 2006, *Fishing rights and small-scale fishers: An evaluation of the rights allocation process and the utilisation of fishing rights in South Africa*, Instituut voor Milieuvraagstukken, Amsterdam.

[254] Jowett B & Barnes J, 1984, *The complete works of Aristotle*, Princeton University Press, Princeton (NJ).

[255] Jung CG, 2014, *The archetypes and the collective unconscious*, 9th Edition, Routledge, London.

[256] Jungk R & Mullert N, 1987, *Future workshops: How to create desirable futures*, Institute for Social Inventions, London.

[257] Kahneman D, 2000, *Evaluation by moments: Past and future*, pp. 693–708 in Kahneman D & Tversky A (Eds), *Choices, values, and frames*, Cambridge University Press, New York (NY).

[258] Kahneman D, 2002, *Maps of bounded rationality: A perspective on intuitive judgment and choice*, Nobel Prize Lecture, **8**, pp. 351–401.

[259] Kahneman D, 2003, *A perspective on judgment and choice: Mapping bounded rationality*, American Psychologist, **58(9)**, pp. 697–720.

[260] Kahneman D & Tversky A, 1972, *Subjective probability: A judgment of representativeness*, Cognitive Psychology, **3(3)**, pp. 430–454.

[261] Kahneman D & Tversky A, 1973, *On the psychology of prediction*, Psychological Review, **80(4)**, pp. 237–251.

[262] Kahneman D & Tversky A, 1979, *Prospect theory: An analysis of decision under risk*, Econometrica, **47(2)**, pp. 263–291.

[263] Kahneman D & Tversky A, 1981, *The simulation heuristic*, (Unpublished) Technical Report, Stanford University, Stanford (CA).

[264] Kant I, 1784, *What is enlightenment?*, pp. 54–60 in Reiss HS (Ed), *Kant's political writings*, Cambridge University Press, Cambridge.

[265] Kant I, 1999, *Critique of pure reason*, 1st Edition, Cambridge University Press, New York (NY).

[266] Kaufmann L, Carter CR & Buhrmann C, 2010, *Debiasing the supplier selection decision: A taxonomy and conceptualization*, International Journal of Physical Distribution and Logistics Management, **40(10)**, pp. 792–821.

[267] Kelley HH, 1967, *Attribution theory in social psychology*, Proceedings of the Nebraska Symposium on Motivation, Lincoln (NE), pp. 192–238.

[268] Kennedy J, 1993, *Debiasing audit judgment with accountability: A framework and experimental results*, Journal of Accounting Research, **31(2)**, pp. 231–245.

[269] Kgatle MS, 2018, *The role of the church in the #FeesMustFall movement in South Africa: Practical theological reflection*, HTS Theological Studies, **74(1)**, pp. 1–8.

[270] King LA & Napa CK, 1998, *What makes a life good?*, Journal of Personality and Social Psychology, **75(1)**, pp. 156–165.

[271] Klauer KC, Musch J & Naumer B, 2000, *On belief bias in syllogistic reasoning*, Psychological Review, **107(4)**, pp. 852–884.

[272] Klayman J, 1995, *Varieties of confirmation bias*, Psychology of Learning and Motivation, **32**, pp. 385–418.

[273] Kotsiantis SB, Zaharakis I & Pintelas P, 2007, *Supervised machine learning: A review of classification techniques*, Emerging Artificial Intelligence Applications in Computer Engineering, **160**, pp. 3–24.

[274] Kramer RM, 2008, *Presidential leadership and group folly: Re-appraising the role of groupthink in the Bay of Pigs decisions*, pp. 230–249 in Ciulla J, Forsyth D & Genovese MA (Eds), *Leadership at the crossroads*, Praeger, Westport (CT).

[275] Kruger J, Wirtz D, Van Boven L & Altermatt TW, 2004, *The effort heuristic*, Journal of Experimental Social Psychology, **40(1)**, pp. 91–98.

[276] Kuran T & Sunstein CR, 1999, *Availability cascades and risk regulation*, Stanford Law Review, **51**, pp. 683–768.

[277] Langa M, Ndelu S, Edwin Y & Vilakazi M, 2017, *#Hashtag: An analysis of the #FeesMustFall movement at South African Universities*, Centre for the Study of Violence and Reconciliation, Cape Town.

[278] Larrick RP, 2004, *Debiasing*, pp. 316–338 in Koehler DJ & Harvey N (Eds), *Blackwell handbook of judgment and decision making*, Wiley Online Library, Hoboken (NJ).

[279] Larrick RP, Morgan JN & Nisbett RE, 1990, *Teaching the use of cost-benefit reasoning in everyday life*, Psychological Science, **1(6)**, pp. 362–370.

[280] Le Bon G, 1897, *The crowd: A study of the popular mind*, Fischer, Frankfurt.

[281] Lecy JD & Beatty KE, 2012, *Structured literature reviews using constrained snowball sampling and citation network analysis*, [Online], [Cited October 2019], Available from http://ssrn.com/abstract=1992601.

[282] Lee Y, 2004, *Student perceptions of problems' structuredness, complexity, situatedness, and information richness and their effects on problem-solving performance*, PhD thesis, Florida State University, Tallahassee (FL).

[283] Lehman DR & Nisbett RE, 1990, *A longitudinal study of the effects of undergraduate training on reasoning*, Developmental Psychology, **26(6)**, pp. 952–960.

[284] Lenney P & Easton G, 2009, *Actors, resources, activities and commitments*, Industrial Marketing Management, **38(5)**, pp. 553–561.

[285] Lerner JS & Tetlock PE, 1999, *Accounting for the effects of accountability*, Psychological Bulletin, **125(2)**, pp. 225–275.

[286] Levin IP & Gaeth GJ, 1988, *How consumers are affected by the framing of attribute information before and after consuming the product*, Journal of Consumer Research, **15(3)**, pp. 374–378.

[287] Levine RA & Campbell DT, 1972, *Ethnocentrism: Theories of conflict, ethnic attitudes and group conflict*, John Wiley & Sons, New York (NY).

[288] Lienert J, Schnetzer F & Ingold K, 2013, *Stakeholder analysis combined with social network analysis provides fine-grained insights into water infrastructure planning processes*, Journal of Environmental Management, **125**, pp. 134–148.

[289] Likert R, 1932, *A technique for the measurement of attitudes*, Archives of Psychology, **140**, pp. 1–55.

[290] Linden M, 2018, *Narrating the 2015 FeesMustFall movement: Explanations contestations and forms of meaning-making in the public sphere*, PhD thesis, University of Pretoria, Pretoria.

[291] Linville PW & Fischer GW, 1998, *Group variability and covariation: Effects on intergroup judgment and behavior*, pp. 123–150 in Sedikides C, Schopler J & Insko CA (Eds), *Intergroup cognition and intergroup behavior*, Lawrence Erlbaum Associates Publishers, Mahwah (NJ).

[292] Lipton M, 1986, *Capitalism and apartheid: South Africa, 1910–1986*, New Africa Books, Cape Town.

[293] Liu JH & Latane B, 1998, *Extremitization of attitudes: Does thought and discussion-induced polarization cumulate?*, Basic and Applied Social Psychology, **20(2)**, pp. 103–110.

[294] Liu Y, Wu L & Li J, 2019, *Peer-to-peer electricity trading in distribution systems of the future*, Electricity Journal, **32(4)**, pp. 2–6.

[295] Loewenstein G, 1999, *A visceral account of addiction*, pp. 235–264 in Elster J & Skog O (Eds), *Getting hooked: Rationality and addiction*, Cambridge University Press, New York (NY).

[296] Longhurst R, 2003, *Semi-structured interviews and focus groups*, pp. 117–132 in Clifford N, Cope M, Gillespie T & French S (Eds), *Key methods in geography*, Sage, London.

[297] Luszcz M, 1989, *Theoretical models of everyday problem solving in adulthood*, Praeger Publishers, Westport (CT).

[298] Makhafola G, 2015, *Students deliver memorandum to ANC*, [Online], [Cited October 2018], Available from https://www.iol.co.za/news/politics/students-deliver-memorandum-to-anc-1934314#.VlRh4tIzfs0.

[299] Malenka DJ, Baron JA, Johansen S, Wahrenberger JW & Ross JM, 1993, *The framing effect of relative and absolute risk*, Journal of General Internal Medicine, **8(10)**, pp. 543–548.

[300] Mannes AE, Larrick RP & Soll JB, 2012, *The social psychology of the wisdom of crowds*, pp. 227–242 in Krueger JI (Ed), *Frontiers of social psychology: Social psychology and decision making*, Psychology Press, New York (NY).

[301] March JG & Simon HA, 1958, *Organizations*, Wiley, New York (NY).

[302] Marhaba M, Wei H, Rui S, Tianyang Z, Hao L, Yao W, Wei L & Yanhong W, 2014, *Relational self versus collective self*, Journal of Cross-Cultural Psychology, **45(6)**, pp. 959–970.

[303] Markovits H & Nantel G, 1989, *The belief-bias effect in the production and evaluation of logical conclusions*, Memory and Cognition, **17(1)**, pp. 11–17.

[304] Markus H & Kunda Z, 1986, *Stability and malleability of the self-concept*, Journal of Personality and Social Psychology, **51(4)**, pp. 858–866.

[305] Markus H & Wurf E, 1987, *The dynamic self-concept: A social psychological perspective*, Annual Review of Psychology, **38(1)**, pp. 299–337.

[306] Masinga L, 2017, *Gauteng Education Department facing financial difficulties, says MEC*, [Online], [Cited October 2018], Available from https://www.iol.co.za/news/south-africa/gauteng/gauteng-education-department-facing-financial-difficulties-says-mec-11660051.

[307] Mason RO & Mitroff II, 1981, *Challenging strategic planning assumptions: Theory, cases, and techniques*, John Wiley & Sons, New York (NY).

[308] McBride DM & Dosher BA, 2002, *A comparison of conscious and automatic memory processes for picture and word stimuli: A process dissociation analysis*, Consciousness and Cognition, **11(3)**, pp. 423–460.

[309] McGrath JE, 1984, *Groups: Interaction and performance*, Prentice-Hall, Englewood Cliffs (NJ).

[310] McGraw KM & Bloomfield J, 1987, *Social influence on group moral decisions: The interactive effects of moral reasoning and sex role orientation*, Journal of Personality and Social Psychology, **53(6)**, pp. 1080.

[311] McKenzie CR, Liersch MJ & Finkelstein SR, 2006, *Recommendations implicit in policy defaults*, Psychological Science, **17(5)**, pp. 414–420.

[312] Meacham JA & Emont NC, 1989, *The interpersonal basis of everyday problem solving*, Praeger, Westport (CT).

[313] Meier BP & Hinsz VB, 2004, *A comparison of human aggression committed by groups and individuals: An interindividual-intergroup discontinuity*, Journal of Experimental Social Psychology, **40(4)**, pp. 551–559.

[314] Merriam-Webster, 2000, *Definition of complex*, [Online], [Cited September 2017], Available from https://www.merriam-webster.com/dictionary/complex.

[315] Milkman KL, Beshears J, Choi JJ, Laibson D & Madrian BC, 2011, *Using implementation intentions prompts to enhance influenza vaccination rates*, Proceedings of the National Academy of Sciences, **108(26)**, pp. 10415–10420.

[316] Milkman KL, Rogers T & Bazerman MH, 2008, *Harnessing our inner angels and demons: What we have learned about want/should conflicts and how that knowledge can help us reduce short-sighted decision making*, Perspectives on Psychological Science, **3(4)**, pp. 324–338.

[317] Miller LS, 1984, *Bias among forensic document examiners: A need for procedural change*, Journal of Police Science and Administration, **12**, pp. 407–410.

[318] Milne S, Orbell S & Sheeran P, 2002, *Combining motivational and volitional interventions to promote exercise participation: Protection motivation theory and implementation intentions*, British Journal of Health Psychology, **7(2)**, pp. 163–184.

[319] Mingers J & Rosenhead J, 2004, *Problem structuring methods in action*, European Journal of Operational Research, **152(3)**, pp. 530–554.

[320] Mingers J & Taylor S, 1992, *The use of soft systems methodology in practice*, Journal of the Operational Research Society, **43(4)**, pp. 321–332.

[321] Miron-Shatz T, 2009, *Evaluating multiepisode events: Boundary conditions for the peak-end rule*, Emotion, **9(2)**, pp. 206–213.

[322] Mischel W, 1999, *Personality coherence and dispositions in a cognitive-affective personality system (CAPS) approach*, pp. 37–60 in Cervone D & Shoda Y (Eds), *The coherence of personality: Social-cognitive bases of consistency, variability, and organization*, Guilford Press, New York (NY).

[323] Miser HJ, 1985, *Handbook of systems analysis: Overview of uses, procedures, applications, and practice*, Elsevier Science Ltd, Amsterdam.

[324] Mitchell DJ, Edward Russo J & Pennington N, 1989, *Back to the future: Temporal perspective in the explanation of events*, Journal of Behavioral Decision Making, **2(1)**, pp. 25–38.

[325] Mitchell RK, Agle BR & Wood DJ, 1997, *Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts*, Academy of Management Review, **22(4)**, pp. 853–886.

[326] Mizuno S & Akao Y, 1978, *Quality function deployment: A company wide quality approach*, JUSE Press, Tokyo.

[327] Mnookin JL, 2008, *The validity of latent fingerprint identification: Confessions of a fingerprinting moderate*, Law, Probability and Risk, **7**, pp. 127–141.

[328] Mokdad AH, Marks JS, Stroup DF & Gerberding JL, 2004, *Actual causes of death in the United States, 2000*, Jama, **291(10)**, pp. 1238–1245.

[329] Moore DA, Oesch JM & Zietsma C, 2007, *What competition? Myopic self-focus in market-entry decisions*, Organization Science, **18(3)**, pp. 440–454.

[330] Moore DA & Healy PJ, 2008, *The trouble with overconfidence*, Psychological Review, **115(2)**, pp. 502–517.

[331] Morgan G, 1980, *Paradigms, metaphors, and puzzle solving in organization theory*, Administrative Science Quarterly, **25(4)**, pp. 605–622.

[332] Morris M, 2018, *South Africa's deficient education system*, [Online], [Cited November 2018], Available from https://www.news24.com/Analysis/south-africas-deficient-education-system-20180507.

[333] Munro I, 1997, *An exploration of three emancipatory themes within OR and systems thinking*, Journal of the Operational Research Society, **48(6)**, pp. 576–584.

[334] Murray ACM, 2019, *An exploration of the potential for blockchain in a transition to a low-carbon energy system in South Africa*, MSc thesis, Stellenbosch University, Stellenbosch.

[335] Murray ACM, 2019, Founder of *Switch*, [Personal Commuminication], Contactable at andrew.murray@switch.org.za.

[336] Mussweiler T & Strack F, 1999, *Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model*, Journal of Experimental Social Psychology, **35(2)**, pp. 136–164.

[337] Mussweiler T, Strack F & Pfeiffer T, 2000, *Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility*, Personality and Social Psychology Bulletin, **26(9)**, pp. 1142–1150.

[338] Myers DG, 1978, *Polarizing effects of social comparison*, Journal of Experimental Social Psychology, **14(6)**, pp. 554–563.

[339] Myers DG & Lamm H, 1976, *The group polarization phenomenon*, Psychological Bulletin, **83(4)**, pp. 602–627.

[340] Naidoo P, Goko C & Cohen M, 2019, *SA has a sovereign debt problem. Its name is Eskom*, [Online], [Cited October 2019], Available from https://www.fin24.com/Economy/sa-has-a-sovereign-debt-problem-its-name-is-eskom-20190724.

[341] NERSA, 2009, *Welcome to the national energy regulator of South Africa*, [Online], [Cited October 2019], Available from http://www.nersa.org.za/.

[342] NEWSTEAD SE, POLLARD P, EVANS JBT & ALLEN JL, 1992, *The source of belief bias effects in syllogistic reasoning*, Cognition, **45(3)**, pp. 257–284.

[343] NGCOBO Z, 2016, *EWN Ziyanda Ngcobo compares #FeesMustFall 2015 and 2016 dynamics*, [Online], [Cited October 2018], Available from http://www.702.co.za/articles/235552/comparing-feesmustfall-2015-and-2016-dynamics.

[344] NICKERSON RS, 1998, *Confirmation bias: A ubiquitous phenomenon in many guises*, Review of General Psychology, **2(2)**, pp. 175–220.

[345] NISBETT RE, 1993, *Rules for reasoning*, Erlbaum, Hillsdale (NJ).

[346] NISBETT RE, KRANTZ DH, JEPSON C & KUNDA Z, 1983, *The use of statistical heuristics in everyday inductive reasoning*, Psychological Review, **90(4)**, pp. 339–363.

[347] NISBETT RE & ROSS L, 1980, *Human inference: Strategies and shortcomings of social judgment*, Prentice-Hall, Englewood Cliffs (NJ).

[348] NKOSI B, 2015, *Student protest brings Wits University to a standstill*, [Online], [Cited October 2018], Available from https://mg.co.za/article/2015-10-14-student-protest-brings-wits-university-to-a-standstill.

[349] NONJINGE G, 2018, *Education for black people in SA is in a terrible state*, [Online], [Cited November 2018], Available from https://www.huffingtonpost.co.za/gugu-nonjinge/the-real-state-of-south-africa-s-education-system_a_23373107/.

[350] NONJINGE G, 2018, *If education is really a priority, we'll have to do things differently*, [Online], [Cited November 2018], Available from https://www.news24.com/Columnists/GuestColumn/if-education-is-really-a-priority-well-have-to-do-things-differently-20180301.

[351] NORTHCRAFT GB & NEALE MA, 1987, *Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions*, Organizational Behavior and Human Decision Processes, **39(1)**, pp. 84–97.

[352] O'DONOGHUE T & RABIN M, 1999, *Doing it now or later*, American Economic Review, **89(1)**, pp. 103–124.

[353] ORMSTON R, SPENCER L, BARNARD M & SNAPE D, 2014, *The foundations of qualitative research*, pp. 1–25 in RICHIE J, LEWIS J, NICHOLLS C & ORMSTON R (EDS), *Qualitative research practice*, Sage Publications, Los Angeles (CA).

[354] PACHUR T, TODD P, GIGERENZER G, SCHOOLER L & GOLDSTEIN DG, 2011, *The recognition heuristic: A review of theory and tests*, Frontiers in Psychology, **2**, pp. 147–167.

[355] PATHER R, 2015, *#FeesMustFall protesters march to Parliament*, [Online], [Cited September 2018], Available from https://mg.co.za/article/2015-11-09-feesmustfall-protesters-march-to-parliament.

[356] PEDREGOSA F, VAROQUAUX G, GRAMFORT A, MICHEL V, THIRION B, GRISEL O, BLONDEL M, PRETTENHOFER P, WEISS R, DUBOURG V, VANDERPLAS J, PASSOS A, COURNAPEAU D, BRUCHER M, PERROT M & DUCHESNAY E, 2011, *Scikit-learn: Machine learning in Python*, Journal of Machine Learning Research, **12**, pp. 2825–2830.

[357] PERROW C, 1972, *Complex organizations: A critical essay*, Scott Foresman and Company, Glenview (IL).

[358] PETERSON GD, CUMMING GS & CARPENTER SR, 2003, *Scenario planning: A tool for conservation in an uncertain world*, Conservation Biology, **17(2)**, pp. 358–366.

[359] Pettigrew TF, 1979, *The ultimate attribution error: Extending Allport's cognitive analysis of prejudice*, Personality and Social Psychology Bulletin, **5(4)**, pp. 461–476.

[360] Pettigrew TF, 1981, *Extending the stereotype concept*, pp. 303–331 in Hamilton D (Ed), *Cognitive processes in stereotyping and intergroup behavior*, Erlbaum, Hillsdale (NJ).

[361] Plomin R, 1994, *Genetics and experience: The interplay between nature and nurture*, Sage Publications, London.

[362] Plous S, 1989, *Thinking the unthinkable: The effects of anchoring on likelihood estimates of nuclear war*, Journal of Applied Social Psychology, **19(1)**, pp. 67–91.

[363] Pohl R, 2004, *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory*, Psychology Press, Hove.

[364] Pollock TG, Rindova VP & Maggitti PG, 2008, *Market watch: Information and availability cascades among the media and investors in the US IPO market*, Academy of Management Journal, **51(2)**, pp. 335–358.

[365] Popper K, 2005, *The logic of scientific discovery*, Routledge, London.

[366] Pouloudi A & Whitley EA, 1997, *Stakeholder identification in inter-organizational systems: Gaining insights for drug use management systems*, European Journal of Information Systems, **6(1)**, pp. 1–14.

[367] PowerX, 2016, *PowerX Africa's energy exchange*, [Online], [Cited October 2019], Available from http://www.powerx.energy/.

[368] Pruitt DG, 1971, *Choice shifts in group discussion: An introductory review*, Journal of Personality and Social Psychology, **20(3)**, pp. 339–360.

[369] Quattrone GA & Jones EE, 1980, *The perception of variability within in-groups and out-groups: Implications for the law of small numbers*, Journal of Personality and Social Psychology, **38(1)**, pp. 141–152.

[370] Quattrone GA, Lawrence CP, Finkel SE & Andrus DC, 1984, *Explorations in anchoring: The effects of prior range, anchor extremity, and suggestive hints*, Unpublished manuscript, Stanford University, Stanford (CA).

[371] Quesada J, Kintsch W & Gomez E, 2002, *A computational theory of complex problem solving using latent semantic analysis*, Proceedings of the 24th Annual Conference of the Cognitive Science Society, Fairfax County (VA), pp. 750–755.

[372] Quintal G, 2015, *What you need to know about #FeesMustFall*, [Online], [Cited November 2018], Available from https://mg.co.za/article/2015-10-19-four-things-you-need-to-know-about-feesmustfall.

[373] Rachel W & Jeevan V, 2019, *University tuition fees hike 'will deter most poorer students' — Poll*, [Online], [Cited September 2019], Available from https://www.theguardian.com/education/2010/nov/18/ipsos-mori-poll-tuition-fees-cuts.

[374] Raune D, MacLeod A & Holmes EA, 2005, *The simulation heuristic and visual imagery in pessimism for future negative events in anxiety*, Clinical Psychology and Psychotherapy, **12(4)**, pp. 313–325.

[375] Reed MI, 1992, *The sociology of organizations: Themes, perspectives and prospects*, Harvester Wheatsheaf, Birmingham.

[376] Reed MS, Graves A, Dandy N, Posthumus H, Hubacek K, Morris J, Prell C, Quinn CH & Stringer LC, 2009, *Who's in and why? A typology of stakeholder analysis methods for natural resource management*, Journal of Environmental Management, **90(5)**, pp. 1933–1949.

[377] Reese EJ, 2012, *Techniques for mitigating cognitive biases in fingerprint identification*, UCLA Law review, **59**, pp. 1252–1290.

[378] Risinger DM, Saks MJ, Thompson WC & Rosenthal R, 2002, *The Daubert/Kumho implications of observer effects in forensic science: Hidden problems of expectation and suggestion*, California Law Review, **90**, pp. 1–56.

[379] Ritchey T, 2011, *Wicked problems — Social messes: Decision support modelling with morphological analysis*, Springer Science & Business Media, New York (NY).

[380] Ritchie T, Skowronski JJ, Hartnett J, Wells B & Walker WR, 2009, *The fading affect bias in the context of emotion activation level, mood, and personal theories of emotion change*, Memory, **17(4)**, pp. 428–444.

[381] Ritov I & Baron J, 1990, *Reluctance to vaccinate: Omission bias and ambiguity*, Journal of Behavioral Decision Making, **3(4)**, pp. 263–277.

[382] Rittel HWJ & Webber MM, 1973, *Dilemmas in a general theory of planning*, Policy Sciences, **4(2)**, pp. 155–169.

[383] Rittel HW & Webber MM, 1973, *Planning problems are wicked*, Polity, **4**, pp. 155–169.

[384] Roberts NC & Bradley RT, 1991, *Stakeholder collaboration and innovation: A study of public policy initiation at the state level*, Journal of Applied Behavioral Science, **27(2)**, pp. 209–227.

[385] Roberts R, Flin R & Cleland J, 2015, *Everything was fine: An analysis of the drill crew's situation awareness on Deepwater Horizon*, Journal of Loss Prevention in the Process Industries, **38**, pp. 87–100.

[386] Robertson LS, 1977, *Car crashes*, Journal of Community Health, **3(2)**, pp. 136–141.

[387] Roth WM & McGinn MK, 1997, *Toward a new perspective on problem solving*, Canadian Journal of Education, **22(1)**, pp. 18–32.

[388] Rothman AJ & Salovey P, 1997, *Shaping perceptions to motivate healthy behavior: The role of message framing*, Psychological Bulletin, **121(1)**, pp. 3–19.

[389] Rothman AJ & Updegraff JA, 2010, *Specifying when and how gain -and loss-framed messages motivate healthy behavior: An integrated approach*, pp. 257–278 in Keren G (Ed), *Perspectives on framing*, Psychology Press, New York (NY).

[390] Rouhana NN & Bar-Tal D, 1998, *Psychological dynamics of intractable ethnonational conflicts: The Israeli–Palestinian case*, American Psychologist, **53(7)**, pp. 761–770.

[391] Rozin P, Millman L & Nemeroff C, 1986, *Operation of the laws of sympathetic magic in disgust and other domains*, Journal of Personality and Social Psychology, **50(4)**, pp. 703–712.

[392] Rozin P & Royzman EB, 2001, *Negativity bias, negativity dominance, and contagion*, Personality and Social Psychology Review, **5(4)**, pp. 296–320.

[393] Russell B, 2013, *History of western philosophy: Collectors edition*, 9th Edition, Routledge, London.

[394] Russell S & Norvig P, 2009, *Artificial intelligence: A modern approach*, 3rd Edition, Pearson Education, London.

[395] Russo J & Schoemaker P, 2016, *Overconfidence*, pp. 1–11 in Augier M & Teece DJ (Eds), *The Palgrave encyclopedia of strategic management*, Palgrave Macmillan, Basingstoke.

[396] SA News, 2019, *Eskom's problems are financial, structural and operational*, [Online], [Cited October 2019], Available from https://businesstech.co.za/news/energy/306202/eskoms-problems-are-financial-structural-and-operational-chairman/.

[397] Al-Saadi H, 2019, *Demystifying ontology and epistemology in research methods*, [Online], [Cited October 2014], Available from https://www.researchgate.net/publication/260244813_Demystifying_Ontology_and_Epistemology_in_Research_Methods.

[398] Salga, 2016, *About us: Welcome*, [Online], [Cited October 2019], Available from https://www.salga.org.za/About%20Us%20W.html.

[399] Salga, 2017, *Status of small scale embedded generation in South African municipalities*, [Online], [Cited October 2019], Available from https://www.salga.org.za/SALGA%20Energy%20Summit%202018/Energy%20Summit%20Web/Document/Status%20of%20Small%20Scale%20Embedded%20Generation.pdf.

[400] Samuel AL, 1959, *Some studies in machine learning using the game of checkers*, IBM Journal of Research and Development, **3(3)**, pp. 210–229.

[401] Sanders GS & Baron RS, 1977, *Is social comparison irrelevant for producing choice shifts?*, Journal of Experimental Social Psychology, **13(4)**, pp. 303–314.

[402] Saribay SA & Andersen SM, 2007, *Are past relationships at the heart of attachment dynamics? What love has to do with it*, Psychological Inquiry, **18(3)**, pp. 183–191.

[403] Schacter DL, 1999, *The seven sins of memory: Insights from psychology and cognitive neuroscience*, American Psychologist, **54(3)**, pp. 182–203.

[404] Schoderbek CG, Schoderbek PP & Kefalas AG, 1980, *Management systems: Conceptual considerations*, Management Education and Development, **7**, pp. 47–48.

[405] Schonfield D & Robertson B, 1966, *Memory storage and aging*, Canadian Journal of Psychology, **20(2)**, pp. 228–236.

[406] Schulz-Hardt S, Frey D, Luthgens C & Moscovici S, 2000, *Biased information search in group decision making*, Journal of Personality and Social Psychology, **78(4)**, pp. 655–669.

[407] Schwarz N, Bless H, Strack F, Klumpp G, Rittenauer-Schatka H & Simons A, 1991, *Ease of retrieval as information: Another look at the availability heuristic*, Journal of Personality and Social psychology, **61(2)**, pp. 195–202.

[408] Sedikides C & Brewer MB, 2015, *Individual self, relational self, collective self*, Psychology Press, New York (NY).

[409] Sedikides C & Gaertner L, 2001, *A homecoming to the individual self*, pp. 7–24 in Sedikides C & Brewer M (Eds), *Individual self, relational self, collective self*, Psychology Press, Philadelphia (PA).

[410] Sedikides C, Gaertner L, Luke MA, O'Mara EM & Gebauer JE, 2013, *A three-tier hierarchy of self-potency: Individual self, relational self, collective self*, Advances in Experimental Social Psychology, **48(1)**, pp. 235–295.

[411] Sedikides C, Gaertner L & O'Mara EM, 2011, *Individual self, relational self, collective self: Hierarchical ordering of the tripartite self*, Psychological Studies, **56(1)**, pp. 98–107.

[412] Sedlmeier P, 1999, *Improving statistical reasoning: Theoretical models and practical implications*, Erlbaum, Mahwah (NJ).

[413] Semple P, 2019, *South Africa's energy sector: On the cusp of a major transition*, [Online], [Cited October 2019], Available from https://www.futuregrowth.co.za/newsroom/south-africa-s-energy-sector-on-the-cusp-of-a-major-transition/.

[414] Serfontein E & De Waal E, 2015, *The corruption bogey in South Africa: Is public education safe?*, South African Journal of Education, **35(1)**, pp. 1–12.

[415] Shafir E & LeBoeuf RA, 2004, *Context and conflict in multiattribute choice*, pp. 341–359 in Koehler DJ & Harver N (Eds), *Blackwell handbook of judgment and decision making*, Blackwell Publishing Ltd, Oxford.

[416] Shepard RN, 1967, *Recognition memory for words, sentences, and pictures*, Journal of Verbal Learning and Verbal Behavior, **6(1)**, pp. 156–163.

[417] Sherif M, Harvey O, White B, Hood W & Sherif C, 1961, *Intergroup conflict and cooperation: The Robbers Cave experiment*, University Book Exchange, Norman (OK).

[418] Sherif M & Sherif CW, 1965, *Ingroup and intergroup relations: Experimental analysis*, pp. 221–266 in Whittaker JO (Ed), *Social psychology: Handbook of basic principles*, Harper Academic, New York (NY).

[419] Siddiqi H, 2016, *Anchoring and adjustment heuristic in option pricing,* Social Science Research Network, [Online], [Cited September 2017], Available from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1636247.

[420] Siegel-Jacobs K & Yates JF, 1996, *Effects of procedural and outcome accountability on judgment quality*, Organizational Behavior and Human Decision Processes, **65(1)**, pp. 1–17.

[421] Simon HA, 1957, *Models of man: Social and rational*, Wiley, New York (NY).

[422] Skues R, 1998, *The first casualty: The war over psychoanalysis and the poverty of historiography*, History of Psychiatry, **9(34)**, pp. 151–177.

[423] Slovic P, 2001, *Cigarette smokers: Rational actors or rational fools?*, pp. 97–124 in Slovic P (Ed), *Smoking: Risk, perception & policy*, Sage Publications, London.

[424] Slovic P, Finucane ML, Peters E & MacGregor DG, 2002, *Rational actors or rational fools: Implications of the affect heuristic for behavioral economics*, Journal of Socio-Economics, **31(4)**, pp. 329–342.

[425] Slovic P, Finucane ML, Peters E & MacGregor DG, 2007, *The affect heuristic*, European Journal of Operational Research, **177(3)**, pp. 1333–1352.

[426] Slovic P, Fischhoff B & Lichtenstein S, 1977, *Behavioral decision theory*, Annual Review of Psychology, **28(1)**, pp. 1–39.

[427] Smith NC, Goldstein DG & Johnson EJ, 2013, *Choice without awareness: Ethical and policy implications of defaults*, Journal of Public Policy and Marketing, **32(2)**, pp. 159–172.

[428] Snape D & Spencer L, 2003, *The foundations of qualitative research*, pp. 1–23 in Richie J & Lewis J (Eds), *Qualitative research practice*, Sage Publications, Los Angeles (CA).

[429] Soll JB, Milkman KL & Payne JW, 2015, *A user's guide to debiasing*, pp. 924–951 in Koehler DJ & Harvey N (Eds), *The Wiley Blackwell handbook of judgement and decision making*, John Wiley & Sons Ltd., Hoboken (NJ).

[430] Sorensen LT & Vidal RVV, 2008, *Evaluating six soft approaches*, Economic Analysis Working Papers, **7(9)**, pp. 1–20.

[431] South African History Online, 2015, *University of Witwatersrand student protests 2015 timeline*, [Online], [Cited October 2018], Available from https://www.sahistory.org.za/article/university-witwatersrand-student-protests-2015-timeline.

[432] Stanovich KE & West RF, 2000, *Individual differences in reasoning: Implications for the rationality debate?*, Behavioral and Brain Sciences, **23(5)**, pp. 645–665.

[433] Starrs JE, 1971, *Ethical obligations of the forensic scientist in the criminal justice system*, Journal of the Association of Official Analytical Chemists, **54(4)**, pp. 906–910.

[434] Stasser G & Titus W, 1985, *Pooling of unshared information in group decision making: Biased information sampling during discussion*, Journal of Personality and Social Psychology, **48(6)**, pp. 1467–1478.

[435] Stone DN & Ziebart DA, 1995, *A model of financial incentive effects in decision making*, Organizational Behavior and Human Decision Processes, **61(3)**, pp. 250–261.

[436] Stoner JAF, 1961, *A comparison of individual and group decisions involving risk*, PhD thesis, Massachusetts Institute of Technology, Cambridge (MA).

[437] Strack F & Mussweiler T, 1997, *Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility*, Journal of Personality and Social Psychology, **73(3)**, pp. 437–446.

[438] Suchman MC, 1995, *Managing legitimacy: Strategic and institutional approaches*, Academy of Management Review, **20(3)**, pp. 571–610.

[439] Suh EM, 2002, *Culture, identity consistency, and subjective well-being*, Journal of Personality and Social Psychology, **83(6)**, pp. 1378–1391.

[440] Sumner WG, 1906, *Folkways — A study of the sociological importance of usages, manners, customs, mores and morals*, Ginn and Company, Boston (MA).

[441] Sutton RS & Barto AG, 1998, *Reinforcement learning: An introduction*, MIT Press, Cambridge (MA).

[442] Swann WB, Stein-Seroussi A & Giesler RB, 1992, *Why people self-verify*, Journal of Personality and Social Psychology, **62(3)**, pp. 392–401.

[443] Switch, 2018, *Renewable energy trading*, [Online], [Cited October 2019], Available from https://switch.org.za/.

[444] Tajfel H & Turner JC, 1979, *An integrative theory of intergroup conflict*, Social Psychology of Intergroup Relations, **33(47)**, pp. 94–109.

[445] Thaler RH & Sunstein CR, 2008, *Nudge: Improving decisions about health, wealth, and happiness*, Yale University Press, New Haven (CT).

[446] The Citizen, 2018, *Fees Must Fall damages cost universities R786m — Report*, [Online], [Cited October 2018], Available from https://citizen.co.za/news/south-africa/1993086/fees-must-fall-damages-cost-universities-r786m-report/.

[447] The Guardian, 2016, *South Africa: Students attack police as protests over tuition fees escalate*, [Online], [Cited November 2018], Available from https://www.theguardian.com/world/2016/oct/04/south-africa-students-attack-police-protests-tuition-fees-escalate.

[448] Thompson SC, 1999, *Illusions of control — How we overestimate our personal influence*, Current Directions in Psychological Science, **8(6)**, pp. 187–190.

[449] Tongco MDC, 2007, *Purposive sampling as a tool for informant selection*, Ethnobotany Research and Applications, **5**, pp. 147–158.

[450] Triandis HC, 1989, *The self and social behavior in differing cultural contexts*, Psychological Review, **96(3)**, pp. 506–520.

[451] Tshwane T, 2018, *Finally: How government plans to fund free education*, [Online], [Cited October 2018], Available from https://mg.co.za/article/2018-02-21-finally-how-government-plans-to-fund-free-education.

[452] Turner JC, 1991, *Social influence*, Brooks/Cole Publishing, Pacific Grove (CA).

[453] Turner JC, 1999, *Some current issues in research on social identity and self-categorization theories*, Social Identity: Context, Commitment, Content, **3(1)**, pp. 6–34.

[454] Turner JC, Hogg MA, Oakes PJ, Reicher SD & Wetherell MS, 1987, *Rediscovering the social group: A self-categorization theory*, Basil Blackwell, Cambridge (MA).

[455] Tversky A & Kahneman D, 1973, *Availability: A heuristic for judging frequency and probability*, Cognitive Psychology, **5(2)**, pp. 207–232.

[456] Tversky A & Kahneman D, 1975, *Judgment under uncertainty: Heuristics and biases*, pp. 141–162 in Vlek C (Ed), *Utility, probability, and human decision making*, Springer, Dordrecht.

[457] Tversky A & Kahneman D, 1983, *Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment*, Psychological Review, **90(4)**, pp. 293–3015.

[458] Ulrich W, 1983, *Critical heuristics of social planning: A new approach to practical philosophy*, John Wiley & Sons, New York (NY).

[459] University of the Witwatersrand, 2018, *Council Committees*, [Online], [Cited October 2018], Available from https://www.wits.ac.za/about-wits/governance/governing-bodies/council-committees/.

[460] University of the Witwatersrand, 2018, *The Council*, [Online], [Cited October 2018], Available from https://www.wits.ac.za/about-wits/governance/governing-bodies/the-council/.

[461] University of the Witwatersrand, 2018, *Vice-Chancellor's Team*, [Online], [Cited October 2018], Available from https://www.wits.ac.za/about-wits/governance/executive-management/.

[462] Valerio MA, Rodriguez N, Winkler P, Lopez J, Dennison M, Liang Y & Turner BJ, 2016, *Comparing two sampling methods to engage hard-to-reach communities in research priority setting*, BMC Medical Research Methodology, **16(1)**, pp. 146–157.

[463] van der Spuy B, 2019, Founder of *Corvus*, [Personal Commuminication], Contactable at vds.brink@corvus.co.za.

[464] van Swol LM, 2007, *Perceived importance of information: The effects of mentioning information, shared information bias, ownership bias, reiteration, and confirmation bias*, Group Processes and Intergroup Relations, **10(2)**, pp. 239–256.

[465] Varvasovszky Z & Brugha R, 2000, *A stakeholder analysis*, Health Policy and Planning, **15(3)**, pp. 338–345.

[466] Vignoles VL, Chryssochoou X & Breakwell GM, 2002, *Sources of distinctiveness: Position, difference and separateness in the identities of Anglican parish priests*, European Journal of Social Psychology, **32(6)**, pp. 761–780.

[467] Vinokur A & Burstein E, 1974, *Effects of partially shared persuasive arguments on group-induced shifts: A group-problem-solving approach*, Journal of Personality and Social Psychology, **29(3)**, pp. 305–315.

[468] Vinokur A & Burnstein E, 1978, *Depolarization of attitudes in groups*, Journal of Personality and Social Psychology, **36(8)**, pp. 872–885.

[469] Von Bertalanffy L, 1972, *The history and status of general systems theory*, Academy of Management Journal, **15(4)**, pp. 407–426.

[470] Vos JFJ & Achterkamp MC, 2006, *Stakeholder identification in innovation projects: Going beyond classification*, European Journal of Innovation Management, **9(2)**, pp. 161–178.

[471] Vul E & Pashler H, 2008, *Measuring the crowd within: Probabilistic representations within individuals*, Psychological Science, **19(7)**, pp. 645–647.

[472] Walker WR & Skowronski JJ, 2009, *The fading affect bias: But what the hell is it for?*, Applied Cognitive Psychology, **23(8)**, pp. 1122–1136.

[473] Walker WR, Skowronski J, Gibbons J, Vogl R & Thompson C, 2003, *On the emotions that accompany autobiographical memories: Dysphoria disrupts the fading affect bias*, Cognition and Emotion, **17(5)**, pp. 703–723.

[474] Wallach MA, Kogan N & Bem DJ, 1962, *Group influence on individual risk taking*, ETS Research Report Series, **62(5)**, pp. 1–32.

[475] Watson SL & Watson WR, 2013, *Critical systems theory for qualitative research methodology*, pp. 111–127 in Dennis B, Carspecken L & Carspecken PF (Eds), *Qualitative research: A reader in philosophy, core concepts, and practice*, Peter Lang Publishing, Pieterlen.

[476] Weaver W, 1948, *Science and complexity*, American Scientist, **36(4)**, pp. 536–544.

[477] Weber M & Henderson AM, 1947, *The theory of economic and social organization*, Oxford University Press, New York (NY).

[478] Weigold MF & Schlenker BR, 1991, *Accountability and risk taking*, Personality and Social Psychology Bulletin, **17(1)**, pp. 25–29.

[479] Weinstein ND, 1977, *Coping with environmental hazards: Reactions to the threat of crime*, Proceedings of the Meeting of the American Psychological Association, San Francisco (CA), pp. 806–820.

[480] Westley F & Vredenburg H, 1991, *Strategic bridging: The collaboration between environmentalists and business in the marketing of green products*, Journal of Applied Behavioral Science, **27(1)**, pp. 65–90.

[481] White RK, 1977, *Misperception in the Arab-Israeli conflict*, Journal of Social Issues, **33(1)**, pp. 190–221.

[482] Whitehouse AJO, Maybery MT & Durkin K, 2006, *The development of the picture-superiority effect*, British Journal of Developmental Psychology, **24(4)**, pp. 767–773.

[483] Wiener N, 1961, *Cybernetics or control and communication in the animal and the machine*, MIT Press, Cambridge (MA).

[484] Wild S, 2016, *Violence escalates at South African universities*, [Online], [Cited November 2018], Available from https://www.nature.com/news/violence-escalates-at-south-african-universities-1.20869.

[485] Wildschut T, Pinter B, Vevea JL, Insko CA & Schopler J, 2003, *Beyond the group mind: A quantitative review of the interindividual-intergroup discontinuity effect*, Psychological Bulletin, **129(5)**, pp. 698–722.

[486] Wilson B, 1990, *Systems: Concepts, methodologies, and applications*, John Wiley & Sons, New York (NY).

[487] Wittenbaum GM & Stasser G, 1996, *Management of information in small groups*, *What's social about social cognition? Research on socially shared cognition in small groups*, Sage Publications, London.

[488] Wood DJ, 1991, *Corporate social performance revisited*, Academy of Management Review, **16(4)**, pp. 691–718.

[489] Wood DJ & Gray B, 1991, *Toward a comprehensive theory of collaboration*, Journal of Applied Behavioral Science, **27(2)**, pp. 139–162.

[490] Wood PK, 1983, *Inquiring systems and problem structure: Implications for cognitive development*, Human Development, **26(5)**, pp. 249–265.

[491] Wood RE, 1986, *Task complexity: Definition of the construct*, Organizational Behavior and Human Decision Processes, **37(1)**, pp. 60–82.

[492] Zajonc RB, 1998, *Emotions*, pp. 591–632 in Gilbert DT, Fiske ST & Linzey G (Eds), *The handbook of social psychology*, McGraw-Hill, New York (NY).

[493] Zajonc RB, 1980, *Feeling and thinking: Preferences need no inferences*, American Psychologist, **35(2)**, pp. 151–175.

APPENDIX A

Functional network labels

This appendix is dedicated to a description of the labels in the functional network presented in Chapter 12, in Table A.1.

TABLE A.1: *A description of the labels in the functional network related to the problem considered in Chapter 12.*

Label	Description
T1.1	Switch's founders may pursue a third-party energy trading business model.
T1.2	As alternative to third-party energy trading, Switch's founders may pursue a wheeling energy trading business model.
T2.1	Swift has to compete with existing third-party traders.
T2.2	Switch's founders require a trading licence in order to operate in the capacity of a third-party energy trader. Additionally, the dynamic pricing strategy they utilise has to be approved. The National Energy Regulator of South Africa is the legal body that adjudicates these affairs.
T2.3	Switch's value offering as third-party energy trader has the potential to alleviate many of the difficulties experienced in Eskom-owned distribution networks.
T2.4	Switch's value offering as third-party energy trader has the potential to alleviate many of the difficulties experienced in municipally-owned distribution networks.
T2.5	Most of the municipally-owned distribution networks in South Africa still use a static electricity trading pricing strategy and are thus considered formally affected by Switch's novel dynamic fee-based software solution.
T2.6	Eskom-owned distribution networks typically still use a static electricity trading pricing strategy and are thus considered formally affected by Switch's novel dynamic fee-based software solution.
T2.7	Switch software solution seems to be of significant utility to grid-tied micro-grids as it has the capacity to aid in the redistribution of electricity within such a setup.
T2.8	If Switch becomes an energy trading software solution provider, the fee structure it proposes will have to be approved by the National Energy Regulator of South Africa.
T2.9	Depending on how the market responds, existing energy traders may lose or gain market share if Switch elects to enter the market as an energy trader. Either way, they will be affected.

TABLE A.1 (continued): *A description of the labels in the functional network related to the problem considered in Chapter 12.*

Label	Description
T2.10	Energy Exchange is currently the only wheeling energy trader in South Africa.
T3.1	PowerX is currently the only third-party energy trader in South Africa.
T3.2	In order to apply for an energy trading licence, Switch must submit a (ready to use) power purchasing agreement contract that will govern the terms of trade between all parties involved.
T3.3	This is simply a *dummy node* to facilitate the presentation of the functional network.
T3.4	If Switch operates as third-party trader in municipally-owned distribution networks, their dynamic pricing strategy has the potential to disrupt the energy sector — Eskom may lose market share.
T3.5	The South African Local Government Association is an autonomous association consisting of the majority of local governments in South Africa (*e.g.* municipalities). They exist to oversee the governance of the respective districts in South Africa and are thus considered formally affected [398].
T3.6	The Association of Municipal Electricity Utilities is an association of regional electricity utilities and commercial companies that strive to improve the working of electricity utilities in South Africa [18]. There is an overlap between this pursuit and Switch's value offering and, as such, they are considered formally affected.
T3.7	A collection of municipalities already identified to be interested in Switch software solution offering.
T3.8	If municipalities start adopting Switch's offering, they will become less dependent upon Eskom for electricity.
T3.9	Switch software solution seems to be of significant utility to grid-tied micro-grid owners as it has the capacity to aid in the redistribution of electricity within such a setup and thereby make this entity less dependent upon Eskom for electricity.
T4.1	A lawyer will be required to formalise the power purchasing agreement.
T4.2	If Switch pursues the option of operating as a third-party trader in Eskom- or municipally-owned distribution networks, independent power producers will be affected. Switch's approach incorporates a dynamic fee structure that determines the price of electricity according to real-time supply and demand information. Consequently, implementing Switch's solution will induce variability in the amount independent power producers receive per mega-watt of electricity.
T4.3	If Eskom or local municipalities grant Switch permission to operate as third-party energy trader or adopts its wheeling energy trading software solution, the price per mega-watt of electricity will vary according to demand scarcity (within certain limits). The lack of a predetermined cost of electricity may make it difficult for power plants to plan for future production.
T4.4	Implementing the dynamic fee pricing strategy associated with Switch's value offering as either third-party trader or software solution provider will alter the amount prosumers receive per mega-watt of electricity they feed back into the grid.
T4.5	If Eskom or municipalities select to employ Switch as a third-party energy trader, consumers in the associated distribution networks will be affected as the implementation of Switch's solution will alter the cost per mega-watt of electricity as a function of supply *versus* demand.

TABLE A.1 (continued): *A description of the labels in the functional network related to the problem considered in Chapter 12.*

Label	Description
T4.6	Switch's approach to determining the price of electricity will most likely alter the general price of electricity, and thereby effect all electricity users (even those that are not associated with Switch).
T5.1	Power-plants are owned by Eskom.
T5.2	Integrating Switch's software solution with the existing smart-meters will require interaction with the smart-meter owner.
T6.1	Eskom owns the smart-meters that are located in Eskom-owned distribution networks.
T6.2	Local municipalities own the smart-meters in municipally-owned distribution networks.
F1.1	Switch may access capital by acquiring a loan.
F1.2	An effective way of raising capital would be to find an investor willing to commit capital in exchange for equity shares.
F1.3	Companies that pursue objectives similar to Switch may be approached as potential partner organisations. This would involve a reallocation of equity shares in exchange for financial and business support.
F2.1	One of a bank's formal objectives is to invest money in entities in the form of repayable loans. Switch's need for capital, formally affects this objective.
F2.2	An equity sale refers to the sale of the common shares of a company in exchange for capital.
F2.3	Switch may elect to raise capital by organising a token sale, otherwise known as an initial coin offering. In essence, this would involve the sale of a digital currency that is linked to the platform Switch utilises for electricity trading. In other words, the value of the digital currency is a function of the rate of Switch's development.
F2.4	Business incubators specialise in guiding newly established companies through the early stages of their development. This includes access to an established platform of potential investors.
F2.5	Eldo is a group of energy management companies that specialise in smart-meter software, energy efficient generation technologies, and energy management services. They are identified as a potential partner organisation.
F3.1	Switch's equity sale will affect venture capitalist firms and their formal purpose of investing in companies.
F3.2	Switch's digital currency offering will affect token buyers and present them with an opportunity to invest.
F3.3	The South African Renewable Energy Business Incubator was established in 2012 in order to stimulate economic development in the renewable energy sector. They assist businesses in the various stages of development and grant them access to an established platform of potential investors.
F3.4	Blockstarters is a blockchain business incubator that attempts to facilitate innovation and resource sharing in the blockchain sector by bringing together like-minded entrepreneurs, providing them with business support and granting them access to an established network of potential investors.
F3.5	LaunchLab is a business incubator situated in Stellenbosch. It offers entrepreneurs excellent networking opportunities with potential investors, and guidance from leaders in the business sector in order to help them launch their business.

TABLE A.1 (continued): *A description of the labels in the functional network related to the problem considered in Chapter 12.*

Label	Description
R1.1	Non-profit organisations with the mandate to support the development of sustainable energy supply technologies are considered formally affected by Switch's business offering.
R1.2	The Department of Energy is responsible for ensuring the country's sustainable energy needs. As Switch's business offering addresses this concern, they are considered formally affected.
R2.1	The International Renewable Energy Agency works toward furthering the development of renewable energy in South Africa.
R2.2	The overall objective of Switch Africa Green is to support a select group of African countries (including South Africa) to transition into an exclusively green economy.
R2.3	GreenCape is a non-profit organisation, situated in the Western Cape, that drives the widespread adoption of economically viable green technology solutions. It works with businesses, investors, academia and government to help unlock the investment and employment potential of green technologies and services.
R2.4	The main objective of the South African Renewable Energy Council is to stimulate economic development in renewable energy technologies. It formally exists as an umbrella body representing several industrial associations that are involved in renewable energy technologies.
R2.5	The Council for Scientific and Industrial Research is a world-class African research and development organisation established in 1945. It is dedicated to conducting research and supporting technological innovations that contribute to the improved quality of life of South African citizens. As Switch's value offering overlaps with these goals, the Council for Scientific and Industrial Research is considered formally affected.
R2.6	The Africa Renewable Energy Initiative is a transformative, Africa-owned and Africa-led inclusive effort to accelerate and scale up the harnessing of the continent's huge renewable energy potential under the mandate of the African Union.
RF1.1	The combination of Switch's financial needs and its renewable energy mandate induces a unique propagation of effect.
RF2.1	Switch may access capital by applying for grants that are targeted at low-carbon and/or renewable energy initiatives.
RF3.1	The South African National Energy Development Institute was established in 2011 under the National Energy Act with the mandate to direct, monitor and conduct energy research and to support technological innovation that promotes the efficient use of energy throughout the South African economy.
RF3.2	The Development Bank of Southern Africa seeks to play a pivotal role in delivering sustainable developmental infrastructure in South Africa and the rest of the African continent.
RF4.1	The Green Finance Desk is an acting partner of the South African National Energy Development Institute. It works with businesses, investors, academia and government to help unlock the investment and employment potential of green technologies and services.
RF4.2	One of the directors of the Development Bank of Southern Africa is a personal acquaintance of one of the founders of Switch.

APPENDIX B

Feedback from a practitioner

This appendix is dedicated to an account of the feedback received from one of Switch's founders, Mr Andrew Murray [335], in respect of the application of the method for stakeholder identification and the decision making checklist to the context of interest in Chapter 12. Notably, what follows is an exact copy of the feedback received from the founder in question. Neither the format, nor the wording has been altered.

Stakeholder identification

- The stakeholder mapping process was useful for getting a picture of how the various stakeholders and decisions fitted together.
- This helped me to consider exactly who needed to be engaged with depending on the decisions made.
- The map was also useful because it illustrated my thought process and helped me see which decisions were underdeveloped and would benefit from additional exploration.
- It was useful because it made me realise where I was making assumptions that might not necessarily be true and because it helped me see which decisions needed further investigation to find more relevant stakeholders.

Decision-making checklist

Anchoring bias and affective failures — This helped me see that with little effort we could include a water-metering and billing solution into our electricity trading solution.

Affective Failures — This helped me because it made me see that we should not neglect the option of acting as a third party trader because the consequences of doing so might be unacceptable, especially if legislation loosens and a third party trading becomes a more attainable option.

Optimistic bias etc. – This made me aware that I was optimistic about the potential for municipalities to want to adopt and use our software even though that might not be the case. It helped me see that there would be value in including them early on in the developmental process so that we ensure that we design the solution with their needs in mind.

Groupthink – This helped us discover that there were issues we might face that we were not discussing. One was the challenge associated with translating power purchasing agreements into smart contracts, which might not be a possibility.

Final remarks

The process overall was truly useful and beneficial and helped me see how the various decisions and stakeholders fitted together. An area that I feel could use further development is in the identification of decisions/options, and stakeholders that I may not have considered. The model is somewhat limited to the knowledge of the person mapping out the process, and that knowledge might not be accurate (*i.e.* my information) — a way to pressure test and validate that knowledge and my assumptions would be useful. The decision making checklist helped me to consider elements that I would likely not have otherwise considered which I'm sure will prove beneficial in the long term sustainability of what we are doing.